The Effects of Risk and Shocks on Non-Farm Enterprise Development in Rural Ethiopia

Bob Rijkers[a] and Måns Söderbom[b]

October 2012

(a) Office of the Chief Economist, Middle East and Northern Africa (MNACE), The World Bank, J5-140, 1818 H Street, N.W., Washington, D.C. 20433, USA. E-mail: brijkers@worldbank.org.

(b) Corresponding author. Department of Economics, University of Gothenburg, P.O. Box 640, SE 405 30 Gothenburg, Sweden. E-mail: mans.soderbom@economics.gu.se. Telephone: +46 (0)31 786 4332. Fax: +46 (0)31 786 1326.

Abstract: This paper assesses the impact of risk and shocks on non-farm enterprise development in rural Ethiopia using a matched household-enterprise-community dataset augmented with precipitation based measures of risk. The likelihood of running a non-farm firm and the returns to running such firms co-vary with agricultural productivity shocks. Diversification into non-farm activities thus provides limited protection against fluctuations in agricultural performance. This may explain why ex-ante risk has no impact on the likelihood of running a non-farm firm. It nonetheless hampers non-farm enterprise development by repressing investment and inducing sorting into less capital-intensive activities.

Key words: risk, rural, firms, Africa, Ethiopia, entrepreneurship, investment-climate

JEL codes: D13, D22, D92, J24, L25, R33

Acknowledgements: We would like to thank the Ethiopian Central Statistical Agency for being able to make extensive use of the survey data, seminar participants at University of Gothenburg, the 2008 CSAE Conference on Economic Development in Africa, the 2008 Oriel College Oxford Joint Academic Forum on Economics and Finance, as well as Magdi Armin, Kathleen Beegle, Elena Bardasi, Luc Christiaensens, Paul Cichello, Mulat Demeke, Mary-Hallward Driemeier, John Giles, Laketch Mikael Imru, Måns Nerman, Pierella Paci, Gbemisola Oseni, Caterina Ruggeri Laderchi, Nicolas van de Sijpe, Jan Svejnar and especially Stefan Dercon, Francis Teal and Josef Loening for useful discussions and comments. Bob Rijkers gratefully acknowledges financial support from the Prins Bernhard Cultuurfonds and the Herbert and Ilse Frankel Memorial Fund at Oriel College, Oxford. Måns Söderbom gratefully acknowledges financial support from Sida. All errors are our own. The views expressed in this paper are those of the authors and do not necessarily reflect those of the World Bank, its Board of Executive Directors, or the countries they represent.

1. INTRODUCTION

Risk is a prominent explanation for the persistence of rural poverty (see e.g. Alderman and Paxson, 1992; Morduch, 1995, 1999; Townsend, 1995; Fafchamps, 1999, 2003; and Dercon, 2004). In the absence of perfect insurance and credit markets risk averse households may choose activity portfolios that yield less volatile, but lower, returns than some alternative portfolio that maximizes expected returns. Risk can discourage investment in fixed capital (e.g. equipment) by incentivizing households to keep assets in liquid form. In addition to these costly ex-ante risk mitigation mechanisms, strategies to cope with short-lived shocks may have lasting negative effects.[i]

A parallel literature stresses the potential of rural non-farm enterprises to catalyze rural growth and structural transformation (see e.g. Barrett et al., 2001; Lanjouw and Lanjouw, 2001; Haggblade et al., 2007; and Bezu et al, 2011). Despite competing theoretical predictions regarding the effects of risk, which are discussed in section 2 below, the nascent literature on the determinants of non-farm enterprise performance (e.g. Deininger et al., 2007, Jin and Deininger, 2009, Rijkers et al., 2010, and Rijkers and Costa, forthcoming) has not examined the impact of risk. This may be considered surprising since the potentially deleterious effects of risk on poverty reduction and growth would appear particularly pertinent in rural areas, which are characterized both by high risk and by highly imperfect markets. Moreover, a lack of non-farm private sector development is synonymous with a lack of investment, and discouraging investment is precisely one of the key mechanisms by which risk is alleged to stunt growth. A major reason for the scarcity of research on the impact of risk on rural non-farm private sector development is that information on rural non-farm enterprises is scant

(Lanjouw and Lanjouw, 2001; World Bank, 2005). Another reason is that econometrically exogenous information on risk is notoriously difficult to collect.

To help redress this lacuna in the literature, this paper augments a novel matched household-enterprise-community dataset with objective precipitation based proxies for risk and shocks to examine the impact of risk and shocks on non-farm enterprise development in the Amhara region in Ethiopia. More specifically, the paper assesses the *ex-ante* impact of risk and the *ex-post* impact of shocks on i) the likelihood that a household runs a non-farm enterprise and ii) firm performance in terms of productivity, investment and long-run capital accumulation.

In doing so, the paper contributes to the literature in a number of ways. To start with, the data enable us to distinguish between the impact of risk and shocks by virtue of containing historical information on predicted agricultural performance over a 12-year period, as well as on self-reported idiosyncratic shocks and coping capacity. This enables us to distinguish the ex-ante impact of being located in a highly insecure environment (risk) from the impact of negative events (shocks), issues which are often conflated in the literature (Dercon, 2008). Moreover, even though the ex-ante effects of risk are often emphasized in theoretical analyses of its impact on investment, the existing literature has predominantly analyzed the effects of shocks (Fafchamps, 2010).

Second, the paper focuses on the implications of risk and shocks on firms' productivity and investment rather than their impact on household-level consumption outcomes. As observed by Dercon (2008), while there is a great deal of evidence on the functioning and formation of mutual insurance mechanisms and coping strategies adopted by households, "*firms have rarely been surveyed with a clear risk perspective, or the data from these surveys have rarely been analysed using risk and its consequences as the focal point*" (pii16).

Third, the paper assesses the determinants of *rural* non-farm enterprise development. While a sizeable body of evidence attests to the importance of a good investment climate for private sector development and growth, most of the literature relies on surveys of relatively large urban-based manufacturing firms (see e.g. Bigsten and Söderbom, 2006). Consequently, it is not clear how the conclusions based on these studies generalize to rural areas, where firms tend to be smaller and enterprise performance is more closely intertwined with household events and agricultural productivity. Our matched enterprise-household-community dataset allows us to examine the impact of such household factors and helps us analyze the relationship between local agricultural performance and non-farm enterprise development.

Fourth and related, it is not clear whether the recent expansion of the non-farm sector in developing countries (Haggblade et al, 2007) is driven by the creation of productive employment opportunities in that sector or by the lack of jobs in the agricultural sector combined with a lack of wage-labor opportunities in urban areas limiting incentives to migrate. By enabling us to document and analyze the returns to running non-farm firms, the data allow us to shed light on this issue.

Ethiopia provides an interesting context to examine the impact of risk on non-farm enterprise development as more than 80% of the Ethiopian population lives in rural areas, where risks are rife (World Bank, 2005). With rapid population growth and increasing pressure on fertile land, diversification into non-farm sector activities features prominently in the government's poverty alleviation strategy. Since poverty is predominantly a rural phenomenon globally (Dercon, 2009, Chen and Ravallion, 2010), the relevance of our study extends beyond the Ethiopian context. Our findings are especially relevant for African countries, because the majority of them remain highly reliant on the rural sector and because Africa' poor growth performance in the past has often been attributed to its high-risk environment (see e.g. Collier and Gunning, 1999).[ii]

The paper is organized as follows. The next section briefly reviews the literature related to our study and enumerates our key hypotheses. The third section describes how we constructed our dataset and how we measure risk and shocks. Participation patterns are analyzed in section 4, while section 5 analyzes the determinants of productivity. The impact of risk on investment and long-run capital accumulation is analyzed in section 6. A final section concludes.

2. RELATED LITERATURE AND HYPOTHESES

The effects of risk and shocks on non-farm entrepreneurship are theoretically ambiguous. On the one hand, in poor rural contexts concerns about food security may induce excessive specialization in farming. If risk causes caution on the consumption side, it could simultaneously reduce both entrepreneurs' willingness to invest and local demand for non-farm products (Fafchamps, 1999; Sen, 1981). Thus, one may postulate a negative correlation between risk and non-farm enterprise activity. On the other hand, consistent with the common view that diversification of activity portfolios is a response to risk (see e.g. Rosenzweig and Binswanger, 1993), one could hypothesize that households living in risky environments are more likely to run a firm. The extent to which non-farm enterprise activity can contribute to income smoothing will both depend on the covariance between agricultural outcomes and non-farm enterprise returns as well as the relative riskiness of each activity. Clearly, all else equal, the stronger the correlation between the returns to running a non-farm firm and the returns to farming, the less effective is diversification.

Studies that have attempted to assess how much of an impact risk and shocks have on activity choice have yielded mixed results. For example, Dercon and Krishnan (1996) study activity portfolio choices in rural Tanzania and find that risk mitigation considerations are not an important driver of activity choice. By contrast, Kurosaki and Fafchamps (2002) show that farmers' observed cropping patterns in rural Pakistan are consistent with their desire to reduce exposure to input price risk. Similarly, shocks can both force people to terminate their business (Mead and Liedholm, 1998) or force them to take up non-farm activities to generate additional income (see e.g. Rose, 2001, Porter, forthcoming, Bezu and Barrett, forthcoming) when other income earning opportunities are lacking.

Risk also affects firm performance through its impact on factor accumulation. In models of investment under risk and irreversibility, risk increases the option value of waiting and therefore reduces the responsiveness of investment to short-run profitability shocks (see e.g. Dixit and Pindyck, 1994). In the long run, there may be a so-called hangover affect working in the opposite direction if irreversibility prevents the firm from selling capital when its marginal revenue product is low (Abel and Eberly, 1999). The sign of the effect - positive or negative - of an increase in risk on long-run capital accumulation in such a setting is therefore ambiguous. The importance of both risk and irreversibility are borne out by several empirical studies (see e.g. Bliss and Stern, 1982, Patillo, 1998, Vargas Hill, 2010, and Dercon and Christiaensen, forthcoming), but to our knowledge, no study has examined the impact of risk on rural non-farm enterprise performance, where irreversibility is a plausible assumption since second-hand capital markets are typically underdeveloped.

Drawing on the literature summarized above, the first main hypothesis we intend to test is that risk and shocks discourage entrepreneurship, while the alternative hypothesis is that they encourage it. The second chief hypothesis is that risk hampers enterprise performance in terms of productivity and investment, while the alternative hypotheses are that risk has no, or possibly even a positive impact on productivity and entrepreneurs' willingness to invest.

3. THE DATA

(a) The Rural Investment Climate Data

The Ethiopian Rural Investment Climate Survey Amhara (RICS-Amhara) was fielded by the World Bank in collaboration with Ethiopia's Central Statistical Agency (CSA) in December 2006 and

January 2007 and yielded a unique matched dataset containing 2909 households, 729 enterprises, and 149 communities,[iii] representative of the Amhara region. The survey contains detailed information on rural households (including non-enterprise owning households), the non-farm enterprises they operate, and the investment climate in the communities in which they are located, as well as subjective information on shocks and coping ability. For the purposes of this survey, a rural nonfarm enterprise was defined as any income generating activity (trade, production, or service) not related to primary production of crops or livestock undertaken either within the household or in any nonhousing units. In addition, any value addition to primary production (i.e processing) was considered to be a rural nonfarm activity. Specific examples of non-farm enterprise activities include brewing, whole sale of agricultural products, retail of both food and non-food products (either in a fixed shop or on markets), running a restaurant or coffee-shop, knitting, making rugs, and running a taxi (either by car or by bike). Note that this set of activities is illustrative but far from exhaustive since the non-farm sector is very heterogeneous.

(b) Measuring Shocks and Risk

To obtain econometrically exogenous proxies for risk and shocks, the RICS-Amhara was augmented with wereda (i.e. district) level indicators of predicted agricultural performance based on rainfall information, using a subsample of the National Oceanic and Atmospheric Administration's Africa Rainfall Estimates Climatology dataset spanning the period from 1995 to 2006. As alluded to in section 2, non-farm entrepreneurs in rural areas are affected by agricultural performance since the bulk of demand derives from customers making a living from agriculture. Rainfall could also induce relative price effects, but the empirical evidence we have at our disposal suggests these are not of first-order importance since prices for products produced by non-farm enterprises are not strongly correlated with agricultural performance (see Table 1).[iv]

<TABLE 1 HERE>

Measuring Shocks. As a proxy for contemporaneous demand shocks, we use the 2006 Water Requirement Satisfaction Index (WRSI), which is plausibly an econometrically exogenous indicator of predicted crop performance based on water availability during the growing season, computed using rainfall information and measured at the wereda (i.e. district) level. WRSI is expressed as a percentage of maximum yield under ideal growing conditions. That predicted performance is measured relative to community-specific maximum yields aids identification as the WRSI measure essentially captures within-community variability in agricultural performance.

In addition to this objective proxy, self-reported reported illness, death and job-loss shocks serve as indicators of idiosyncratic household-level shocks. A household is considered to have suffered an illness (death/job-loss shock) if at least one household member was reported to suffer an illness (death/job loss) during the year preceding the survey. Finally, a household's risk-coping capacity is proxied by a dummy variable indicating whether the household considers itself capable of raising 100 Birr in case of an emergency. This variable also serves as a crude indicator for a household's access to finance. Since these self-reported measures are subjective, they have to be interpreted with caution.

Measuring risk. To measure ex-ante risk, we use the standard deviation of WRSI over the period from 1995 up until and including 2005. A key strength of our measure of risk is that it arguably captures a salient source of risk facing Ethiopian farmers. In addition, it is objective and plausibly econometrically exogenous, while most previous literature has relied on potentially endogenous and often subjective measures of risk.[v]

4. PARTICIPATION

(a) Rural life is risky

Table 2 presents summary statistics, distinguishing households with and those without an enterprise.[vi] In Amhara, 23% of all households own a non-farm enterprise.[vii] Enterprise owning households on average derive about a third of their income from running a non-farm enterprise, but are not significantly better off than households which do not run a non-farm firm. The log of average annual household expenditure per capita is 7.3, or about $163 (1480 Birr) in levels, which is a stark reminder of how poor the households in our sample are. While there is substantial variation across households, even the relatively best off households are not well off. For example, at the 95th percentile, the average log average annual household expenditure is about 8.3, or about $456 (4146 Birr).

<TABLE 2 HERE>

Prima facie, there is no strong link between ex-ante risk, vulnerability to shocks and enterprise ownership, even though life in rural areas is clearly risky. The standard deviation of predicted agricultural performance is about 7.6% for enterprise owning households and 7.4% for non-enterprise owning households. Shocks are prevalent; 24% of enterprise owning households reported having experienced an illness shock, 4% reported having experienced a death shock and 2% reported having suffered a job loss shock in the year preceding the survey. The incidence of shocks reported by non-enterprise owning households was very similar as 19% reported an illness shock, 2% a death shock and 4% a job loss shock. While the proportion of household reporting job loss shocks may seem low in absolute terms, it is high when one considers that merely 10% of all working-age individuals have worked for a wage at some point during the year.

The descriptive statistics furthermore suggest that enterprise owning households are more likely to be headed by women, individuals who are divorced and migrants, the latter defined here as those who are living in communities in which they were not born (these are local, not international migrants). Moreover, entrepreneurs are more likely to live in rural towns, closer to markets and roads, as well as more likely to live in communities where financial institutions are present. On the other hand, on average, enterprise owning household are neither headed by older or better educated individuals, nor are they better able to access emergency credit if necessary.

(b) Modelling Participation

To assess the impact of shocks and risk on the propensity to run a firm, we estimate a probit model for participation, P, using as explanatory variables household and community characteristics, H and C respectively, risk, U, and indicators of shocks and coping capacity, S:

$$\Pr(P = 1) = \Phi(\pi_H H + \pi_C C + \pi_U U + \pi_S S) \qquad (1)$$

where π_i, $i \in \{H,C,U,S\}$, indicates a vector of parameters to be estimated. The key null hypothesis of interest is that diversification into non-farm activities is not motivated by the desire to mitigate risk i.e., that π_U=0. In addition, we assess whether non-farm enterprise activity responds to shocks by testing whether π_S=0

Omitted variable bias and endogenous placement are two key threats to identification. Since our objective risk and shocks proxies are measured at the community level, being able to control for potentially confounding differences between locations is important. Our data enable us to control for a rich set of objective community characteristics by virtue of containing detailed information on the local investment climate, which should reduce omitted variable bias. To assess the importance of endogenous placement we present regressions where we control for whether or not the household migrated into the community, and assess whether risk has a disproportionate impact on the decision of such families to run a non-farm enterprise. If aspiring entrepreneurs choose to locate in less risky areas to set up firms, one would expect migrant households' propensity to set up non-farm firms to be more strongly correlated with risk than that of non-migrant households, ceteris paribus.

(c) Results

<TABLE 3 HERE>

Table 3 shows estimates from probit models of the likelihood that a household has at least one non-farm firm.[viii] The baseline specification presented in column 1 includes as regressors basic household demographics, characteristics of the household head, location dummies, distance variables, contemporaneous WRSI, and a measure of ex-ante risk, the standard deviation of WRSI over the period 1995-2005. The second specification, which is our preferred specification, adds further community controls for the opportunity cost of labour, the quality of infrastructure and access to credit. While these variables are potentially endogenous to participation, their inclusion serves to reduce potential omitted variable bias. Column 3 adds subjective idiosyncratic shocks variables and a dummy for whether or not the household would be able to generate 100 Birr in case of an emergency, referred to as emergency credit. The final column adds an interaction term between our proxy for risk and being a migrant household to test for the importance of endogenous placement.

The main finding is that non-farm entrepreneurship is not correlated with risk. The finding is robust to including additional controls for community characteristics (column 2) and idiosyncratic shocks and coping capacity (column 3), which are not significantly correlated with participation. The results presented in column 4 furthermore suggest that endogenous placement does not play a major role. While migrant households are much more likely to run a non-farm firm (perhaps because they lack other income earning options) ceteris paribus, their decision to run non-farm firms does not seem correlated with risk, as is indicated by the fact that the coefficient on the interaction term between the standard deviation of WRSI and being a migrant household is close to zero and statistically insignificant.[ix] Overall, these findings suggest that that risk mitigation considerations cannot explain the decision to run a non-farm firm and provide empirical evidence for the proposition that households' activity portfolio choice *cannot be explained by their behaviour towards risk as is usually suggested* (Dercon and Krishnan., 1996, p.850).

By contrast, predicted agricultural performance is positively correlated with participation at the 10% level. The effect of better agricultural performance is significant and non-negligible; a one percent increase in predicted agricultural output is associated with an increase in the participation rate of 0.9%.

The results furthermore suggest that entrepreneurship is more important for women than for men. Female-headed households are much more likely to run a firm and participation is also positively correlated with the number of women in the household. A possible explanation for this finding is that cultural preferences inhibit women from working on the farm. For instance, many households in Amhara believe that the harvest will be bad if women work on the farm (Zwede and Associates, 2002; and Bardasi and Getahun, 2007). Moreover, the results underscore the importance of location; households located in a rural town are some 21 to 24 percentage points more likely to engage in non-farm enterprise activities. Households living further away from food markets are significantly less likely to run a non-farm firm. There is an inverse U relationship between entrepreneurship and the age and education of the household head.[x] Entrepreneurship is further more prevalent amongst divorcees, and significantly higher in areas with better access to electricity. Participation rates are also correlated with the ability to use land as collateral when requesting a loan from a financial institution.

5. PERFORMANCE

(a) Enterprises are small and sales are low

Tables 4 summarizes key enterprise characteristics. Enterprises are very small, highly labor-intensive and not very productive.[xi] More than three quarters of the enterprises in our sample employ only one worker and only three firms employ more than 5 workers. The mean of log annual sales is 7.38, or about $173 (approximately 1600 Birr) in levels. There is a lot of heterogeneity across firms, as is evidenced by the high standard deviation of sales. On average, the returns to running a non-farm enterprise seem lower than the returns to having a wage job. The average of log value added per day worked in a non-farm enterprise is 1.22, equivalent to roughly 3 Birr or about $0.37. By comparison, male agricultural workers typically receive a daily wage per day of 8 Birr, while women receive 7 Birr per day.

<TABLE 4 HERE>

The low average output of non-farm firms is in part due to low capital intensity and material inputs usage; the average log value of the capital stock is 5.10, which corresponds to about $18 (164 Birr), in levels, while the mean log of material inputs is 5.76, corresponding to about $35 (317 Birr).

Turning to the sectoral composition of the non-farm sector, most (56%) of the non-farm enterprises in the sample undertake manufacturing activities. In addition, just under a third (29%) of non-farm enterprises engage in trade and the remainder provide services, e.g. operate a hotel or run a restaurant.

(b) Modelling Strategy

In order to describe and explain productivity differentials across different types of enterprises and localities, we use a simple augmented Cobb-Douglas production function framework in which output, Y, is modelled as a function of capital, K, labour inputs, L, material inputs, M, and TFP, which is in turn modelled to be a function of the characteristics of the entrepreneur, E, and the firm, F, such as its sectoral affiliation, community characteristics, C, risk, U, and subjective shocks and coping capacity, S. We try to correct for possible selection bias in the production function, an endogeneity problem that has received surprisingly little attention in the literature, (Ackerberg et al., 2007), by including a Heckman selection correction term θ (the inverse Mills ratio) based on the extended participation probit regression presented in column 2 of Table 3.[xii] Thus, the exclusion restrictions on which this selection correction relies are household demographics and marital status, which are assumed to affect activity portfolio choice through their impact on the opportunity cost of labour (and capital), but not the returns to different activities conditional on choosing them.[xiii] Selection bias may arise because the amount of time spent running a non-farm firm may be correlated with entrepreneurial ability; those running a non-farm enterprise might be better at it than those who do not, causing a spurious correlation between inputs usage and TFP.[xiv] Alternatively, those running an enterprise might simply lack alternative options, perhaps because they are not very able. A priori, it is thus difficult to sign the selection bias. Augmenting the production function with this selection correction term yields the following estimable equation:

$$\ln Y = \beta_K \ln K + \beta_L \ln L + \beta_M \ln M + \beta_E \ln E + \beta_F F + \beta_C C + \beta_U U + \beta_S S + \delta\theta + \varepsilon \qquad (2)$$

where ε is a zero-mean normally distributed error term, assumed to be uncorrelated with the regressors in the model. The key null hypotheses of interest are that risk and shocks do not impact firm performance, i.e. that: $\beta_U = 0$ and $\beta_S = 0$.

We estimate this production function by means of OLS. A well-documented concern in the literature on production function estimation is that the resulting parameter estimates will be biased if the residual is correlated with inputs. In principle, one might be able to solve this problem using instrumental variables, provided good instruments are available. Unfortunately, in our dataset such instruments are not available.[xv] However, a large set of control variables may go a long way towards controlling for unobserved productivity. Söderbom and Teal (2004), for example, use a panel of predominantly small manufacturing firms from Ghana to show that instrumental variable estimates of production function parameters are very similar to their OLS counterparts. Furthermore, in the rural context where there is limited economic integration, it would seem plausible to assume that factor prices vary quite a lot across villages (the low correlations between product-specific prices reported in Table 1 are consistent with this argument), which would generate exogenous variation in the factor inputs. Moreover, firms do not invest or change their labor input very much, despite facing frequent shocks. For these reasons, endogeneity bias may be limited.

To assess whether this is indeed the case, we adopt a proxy variable approach and estimate the production function with and without our precipitation based measures of predicted agricultural performance and risk, which serve as crude proxies for local demand. If endogeneity of inputs is a major concern, one would expect the inclusion of these measures to lead to very different parameter estimates on factor inputs. By contrast, if such parameter estimates do not change very much, endogeneity bias is likely more limited.

(c) Results

Results from the production function regressions are shown in Table 5. Column 1 presents the "bare" specification, using factors of production, characteristics of the manager, location dummies, distance measures, firm age and sub-sector dummies as explanatory variables. In column 2, which presents our baseline specification, selection bias is controlled for by means of the Heckman procedure. In addition, precipitation-based measures of ex-ante risk and local demand are added. Column 3, which is our preferred specification, adds community controls for infrastructure, local wages rates, access to credit and competition. Though they are potentially endogenous to performance, the inclusion of these variables should reduce omitted variable bias. Column 4 includes subjective shocks measures and an indicator of self-reported coping capacity.

<TABLE 5 HERE>

The models shown in Table 5 explain a substantial part of the variation in sales. The factors of production – labour days, capital and materials - are all significant. While the assumption of constant returns to scale is never rejected by the data, our production function estimates hint at the existence of very mildly increasing returns to scale since the sum of the coefficients on capital, labour and material inputs is always larger than 1. The estimated coefficient on labour is about 0.7, which though high is in line with the literature on microenterprises (see e.g. Deininger et al. (2007) and Jin and Deininger (2009) for comparable estimates for Sri Lankan and Tanzanian rural microenterprises). The coefficients on capital and inputs are rather modest.

The results further suggest modest or no selection bias: the coefficient estimates on Heckman's lambda are very small and statistically insignificant except in column 2 where it is marginally significant. It is possible that the insignificance of Heckman's lambda is due to weak instruments. Recall that the assumption behind the exclusion restrictions is that household demographics and marital status impact participation only, and not performance directly. To investigate whether weak instruments pose a problem, we revisit the participation probit results reported in Table 3. The fact that divorced/separated, and the share of adult women in the household, are statistically significant at the 1% level in the participation regressions (see Table 3) suggests that the instruments are sufficiently strong.

To investigate whether endogeneity bias is a problem we compare the results in columns 1 and 2 in Table 5. Recall that one would anticipate precipitation-based measures of agricultural performance to be strongly correlated with inputs usage if endogeneity of inputs were a major concern. However, the coefficient estimates in columns 1 and 2 are very similar, even though we add controls for unobserved demand and risk in column 2. Moreover, the contemporaneous WRSI measure enters strongly significantly, which is consistent with our hypothesis that this variable is a good proxy for unobserved demand. These findings thus suggest that the likely impact of endogeneity of inputs bias is small. The results are robust to including additional community and investment climate controls (column 3) and self-reported shocks (column 4). The impact of risk and other shocks on enterprise performance is statistically negligible.

The positive and significant coefficient on contemporaneous WRSI is a key finding of this paper since it implies that (the prospect of) a good harvest raises productivity among non-farm enterprises, probably because of higher local demand. An alternative explanation would be that rainfall induces price effects. Recall, however, that the results presented in Table 1 suggest that such price effects are not large. The high covariance between non-farm enterprise sales and agricultural outcomes is consistent with the finding that risk mitigation does not explain non-farm enterprise participation since it indicates that diversification into non-farm enterprise activity is of limited use when attempting to mitigate weather-risk.

Turning to other results of interest, male-operated enterprises are more productive than enterprises managed by women and firms located in rural towns are more productive than other firms. The relationship between enterprise profitability and the education of the manager is convex.[xvi] The manager's age and its square are neither individually nor jointly significantly correlated with higher productivity. The coefficients on the industry dummies suggest that manufacturing activities are among the least productive activities while trading activities, such as wholesale and retail, are very productive. The high profitability of trading activities could reflect arbitrage opportunities due to limited economic integration, which may also explain why enterprises engaging in transport activities are very profitable.

Finally, it is worth pausing to consider the implied marginal productivity of capital. The estimated median marginal productivity of capital is 77%. Our estimates are in line with the very high returns to capital in microenterprises documented by McKenzie and Woodruff (2006) and de Mel et al. (2007), who estimate that the average annual returns to capital in Sri Lankan microenterprises are at least 68% per year.

6. INVESTMENT AND LONG-RUN CAPITAL ACCUMULATION

(a) Modelling Strategy

Despite very high marginal returns to capital, only 18% of firms in our sample record any investment in fixed capital. In addition, the firms that do, typically invest only very small amounts and average capital stocks are very low.[xvii] To examine to what extent risk accounts for such low levels of capital accumulation, we estimate a simple investment probit. Our most general investment model uses as explanatory variables initial conditions, *B*, enterprise characteristics, *E*, community characteristics, *C,* risk, *U,* covariate and individual shocks and coping capacity, *S* and the inverse Mills ratio θ to correct for selection bias:[xviii]

$$\Pr(I > 0) = \Phi(\tau_B B + \tau_E E + \tau_C C + \tau_U U + \tau_S S + \tau_\theta \theta) \quad (3)$$

where *I* denotes investment.

The investment probit identifies the short-term impact of risk on investment. To examine the impact of risk on long-run capital accumulation we model the capital stock and the capital labor ratio as a function of the same set of explanatory variables:

$$\ln K = k_B B + k_E E + k_C C + k_U U + k_S S + k_\theta \theta + w \quad (4)$$

$$\ln K/L = c_B B + c_E E + c_C C + c_U U + c_S S + c_\theta \theta + z \quad (5)$$

where w and z are exogenous, zero-mean, normally distributed error terms.

We present regressions without and with controls for initial conditions. In addition, in our models of the long-run capital stock and capital intensity we present regressions that do not control for sector since activity choice and capital intensity are closely intertwined.

(b) Results

(i) Investment

Table 6 shows results from probit regressions modelling the likelihood of buying equipment (investment). The baseline specification, presented in column 1, controls for firm age, sectoral affiliation, characteristics of the manager, location dummies, distance measures, risk, as well as contemporaneous predicted rainfall. In the second column additional community and investment climate controls are added. The third column adds subjective shocks measures and a dummy indicating whether or not the household is able to generate 100 Birr in case of an emergency. In the final column, labor usage and capital at startup, proxied by the answer to the question "how much money did you pay to set up the firm?", are added to control for initial conditions.[xix]

<TABLE 6 HERE>

Our principal result is a clear and very strong negative relationship between risk and investment, consistent with conventional models of investment under risk. By contrast, shocks are not statistically significant predictors of investment in any of the specifications. In the risky rural context, where insurance is imperfect and access to credit is limited, precautionary motives provide incentives for households to store wealth in the form of liquid savings rather than make irreversible investments in illiquid assets. This suggests that market failures in insurance markets may hamper enterprise performance by suppressing investment.

Indeed, the household's ability to raise credit in an emergency is strongly positively correlated with investment (see columns 3 and 4). While this variable is likely endogenous, this suggests that households who are better capable of mitigating shocks are more willing to assume risk. This finding might also be interpreted as indicating that access to credit is an important determinant of investment. Yet, the coefficient on the financial institution dummy is anomalously negative and significant at the 10% level. Overall then, the evidence for the idea that access to credit is a major impediment to investment is mixed and not particularly strong.

The initial capital stock of the enterprise (column 4) is strongly positively associated with the probability of investment, indicating that larger firms are more likely to invest. Older firms are also more likely to invest. This could be because, as time goes on, upgrading the capital stock becomes more important. Alternatively, it could be that young firms are faced with higher risk regarding the prospects of the enterprise, which may lead to caution on the investment side. The propensity to invest also varies significantly across sectors, and male managers seem more likely to invest than female ones. In addition, firms that face more competitors are somewhat more likely to invest. On the other hand, there is no evidence that firms located in remote areas invest less, nor for the idea that investment behaviour varies with the age or education of the manager. Selection bias does not appear to play a major role, as is evidenced by the small and insignificant coefficients on the inverse Mills ratio.

(ii) Long Run Capital Accumulation

Although risk is negatively correlated with investment, its impact on long-run capital accumulation is theoretically ambiguous (see e.g. Abel and Eberly, 1999, for an analysis of the relationship between risk and long-run capital accumulation based on a theoretical model in which investment is fully irreversible). Table 7 presents models of the capital stock. Columns 1 through 4 mimic the specifications for investment, except that they do not include controls for sector. Instead we add separate specifications where we replicate specification 3 and control for sector choice (column 5), and both sector choice and initial conditions (column 6).

<TABLE 7 HERE>

The determinants of the capital stock and the determinants of investment are similar. Risk is strongly negatively correlated with investment and this effect is robust to controlling for community characteristics (column 2), as well as shocks and the household's ability to raise emergency credit (column 3). However, the negative association between risk and capital usage becomes insignificant when we include controls for initial conditions (column 4) sector (column 5) or both (column 6). Thus, it appears that risk induces sorting into less capital intensive activities: once this sorting effect is accounted for, the negative association between risk and capital vanishes.

Households which have the ability to generate 100 Birr in case of an emergency also run firms with higher capital stocks, ceteris paribus. Though supportive of the hypothesis that households that are better able to cope with risk are more likely to invest and accumulate capital, this finding may reflect reverse causation; having more collateral may make it easier to obtain credit and invest. Self-reported illness, death and job-loss shocks are not correlated with the long-run capital stock.

One might be concerned that the capital stock is merely proxying for the determinants of firm-size, rather than reflecting capital intensity. As a robustness check, table 8 presents specifications that are identical to those presented in table 7, but now using the capital labour ratio as the dependent variable. The results are very similar to those obtained using the log of the capital stock as the dependent variable; risk is negatively correlated with capital intensity, but this effect becomes statistically negligible once initial conditions or sector are controlled for.

<TABLE 8 HERE>

A related concern is that the productivity and investment regressions presented in sections 4 and 5 obfuscate an impact of risk since they all control for sector. To alleviate this concern we have also examined the impact of risk on investment and productivity not controlling for sector. The results which are not presented to conserve space but available from the authors upon request, remain qualitatively unchanged.

As a final robustness check, Table 9 presents quantile regressions which help us examine how the relationship between risk and capital accumulation varies across the distribution of the capital stock using the same specification as presented in column 1 of table 7. While not always statistically significant, the coefficient estimate associated with weather volatility is always negative. Interestingly, the negative association is substantially stronger at higher quantiles of the capital distribution, where it is also consistently statistically significant, albeit at the 10% level only. The results thus suggest that the relationship between risk and capital accumulation is most pronounced for firms using relatively large amounts of capital.

To sum up, our results suggest that risk represses investment and induces sorting into less capital intensive activities.

7. CONCLUSION

The proposition that risk perpetuates rural poverty by repressing investment and inducing households to forsake potentially more profitable income-earning opportunities to have a more stable income-stream features prominently in current development debates. A parallel and growing literature emphasizes the potential of small rural non-farm enterprises to catalyze rural poverty reduction and growth. Despite competing theoretical predictions about the impact of risk on rural non-farm enterprise development, the empirical evidence on how risk and shocks affect rural non-farm enterprise performance remains sparse. This is surprising since rural life is rife with risks and since the repercussions of high risk are likely most severe when insurance and credit markets are imperfect or incomplete, as is often the case in rural areas. The lack of empirical evidence on the relationship between risk and non-farm enterprise development reflects data-limitations, as information on both rural non-farm enterprises and risk is sparse.

To help redress this lacuna in the literature, this paper augments a novel matched household-enterprise-community dataset from the Amhara region in Ethiopia with precipitation based measures of risk to examine the impact of risk and shocks on the decision to become an entrepreneur and on firm performance in terms of productivity, investment and long-run capital accumulation. We would argue that our risk and shocks proxies are econometrically exogenous and capture a salient source of risk faced by rural households.

Participation in non-farm enterprise activity is positively correlated with predicted contemporaneous agricultural performance. By contrast, risk does not affect the decision to diversify per se. A plausible explanation for these findings is that diversification into non-farm enterprise activity is not an effective ex-ante means of mitigating fluctuations in agricultural performance since non-farm enterprise sales covary strongly with predicted agricultural performance. However, ex-post the returns to running a non-farm firm are higher when agricultural performance is stronger, which may explain why participation rates are responsive to shocks.

This does not mean, however, that risk does not act as a drag on rural non-farm private sector development. Risk is negatively correlated with investment and appears to induce sorting into less capital intensive activities: it helps explain why investment rates are low even though the implied returns to capital are very high. The negative association between risk and capital accumulation is the most pronounced for firms that use the most capital. Alternative, potentially complementary, explanations for this paradoxical coexistence of high marginal returns and low investment are high discount factors (see e.g. Bluffstone and Yesuf, 2007) and prohibitively high transaction costs associated with supplying markets outside the perimeter of one's locality. The data lend some support to the latter explanation as non-farm enterprise activity is concentrated in rural towns, where the returns to running a non-farm firm are significantly higher. In conjunction with the fact that most enterprises are very small and enterprise productivity is highly dispersed, this suggests that market fragmentation is an important explanation of the poor performance of Ethiopia's rural non-farm sector.

Although they do not appear very suitable instruments for mitigating agricultural output risk, non-farm enterprises nonetheless fulfill an important safety net function as they generate income for those with limited income earning opportunities, such as widows, women and migrants. Although the returns to running a non-farm firm are low on average, and much lower than prevailing local agricultural wage rates, participation rates are considerable, suggesting that many of the non-farm entrepreneurs lack better options. In other words, despite high heterogeneity in firm performance, the non-farm sector is not as yet a major engine of productive employment creation.

While our findings are suggestive of the ways in which risk and shocks affect non-farm enterprise development, the analysis is limited to very specific risks. Collecting data on other types of risk, such as those related to governance, transactions and labor exchange, and tracing enterprises and households over time would facilitate better identification of the impact of risk on risk and permit a richer characterization of non-farm enterprise dynamics and how they are impacted by risk and shocks.

REFERENCES

Abel, A. B. and Eberly, J. (1999). "The effects of irreversibility and uncertainty on capital accumulation," *Journal of Monetary Economics* 44(3): 339-377.

Ackerberg, D., Benkard, Lanier, C., Berry, S. and Pakes, A. (2007) "Econometric Tools for Analyzing Market Outcomes" *Forthcoming chapter in Handbook of Econometrics*, Vol. 6A. Elsevier, North Holland.

Alderman, H. and Paxson, C. (1992). "Do the poor insure? A synthesis of the literature on risk and consumption in developing countries," World Bank Policy Research Working Paper No. 1008.

Andersson, A., and Jirström, M. (2007) "Rural-Urban Linkages in Amhara" Department of Social and Economic Geography. Lund University: Lund. Processed draft report.

Bardasi, E. and Getahun, A. (2007) "Gender and Entrepreneurship in Ethiopia" Background paper to the Ethiopia Investment Climate Assessment Washington DC: The World Bank.

Barrett, C. Reardon, T. and Webb, P. (2001) "Non-farm Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics and Policy Implications" *Food Policy* 26(4): 315-331

Bezu, S, Barrett, C., and Holden, T. (2011) "Does non-farm economy offer pathways for upward mobility? Evidence from a panel dataset from Ethiopia" *Cornell University Working Paper.*

Bezu, S. And Barrett, C.(forthcoming) "Employment dynamics in the rural nonfarm sector in Ethiopia: Do the poor have time on their side?" *Journal of Development Studies.*

Bigsten, A. and M. Söderbom (2006). "What Have We Learned from a Decade of Manufacturing Enterprise Surveys in Africa?" *World Bank Research Observer* 21(2): 241-265.

Bliss, C. and Stern, N. (1982) *Palanpur: The Economy of an Indian Village,* Oxford, UK: Clarendon Press.

Bluffstone, R. and Yesuf, M. (2007) “Risk aversion in low income countries: Experimental evidence from Ethiopia” IFPRI Discussion Paper 715.

Chen, S. and Ravallion, M. (2010) "The Developing World Is Poorer Than We Thought, but No Less Successful in the Fight Against Poverty," *The Quarterly Journal of Economics* 125(4): 1577-1625.

Collier, P. and Gunning, J. (1999) "Explaining African Economic Performance" *Journal of Economic Literature* 37(1): 64-111.

de Janvry, A, Finan, F., Sadoulet, E. And Vakis, R. (2006) “Can conditional cash transfers serve as safety nets to keep children at school and out of the labor market?” *Journal of Development Economics* 79(2):349-373.

de Mel, S., McKenzie, D. and Woodruff, C. (2007) “Returns to capital in microenterprises: evidence from a field experiment” World Bank Policy Research Working Paper No.4230.

Deininger, K., Songqing, J. and Sur, M. (2007) "Sri Lanka's Rural Non-Farm Economy: Removing Constraints to Pro-Poor Growth" *World Development* 35(12): 2056- 2078.

Dercon, S. (2004) *Insurance against poverty* Oxford, UK: Oxford University Press.

Dercon, S. (2008) “Fate and Fear: Risk and Its Consequences in Africa” *Journal of African Economies* 17(2): ii97-ii127.

Dercon, S. (2009). "Rural Poverty: Old Challenges in New Contexts," *World Bank Research Observer*, 24(1):1-28.

Dercon, S. and Christiaensen, L. (forthcoming) “Consumption risk, Technology Adoption and Poverty Traps: Evidence from Ethiopia” *Journal of Development Economics.*

Dercon, S. and Krishnan, P. (1996) "Income Portfolios in Rural Ethiopia and Tanzania: choices and constraints" *Journal of Development Studies* 32(6): 850-75.

Dercon, S. and Krishnan, P. (2000) "In Sickness and in Health: Risk-sharing within Households in Ethiopia", *Journal of Political Economy* 108 (4): 688–727.

Dixit, A. and Pindyck, R. (1994) *Investment Under Uncertainty,* Princeton, NJ: Princeton University Press.

Dollar, D., Hallward-Driemeier, M. and Mengistae, T. (2005) "Investment Climate and Firm Performance in Developing Economies" *Economic Development and Cultural Change* 54(1): 1-31.

Fafchamps, M. (1999) "Rural Poverty, Risk, and Development", FAO, Economic and Social Development Paper No. 144.

Fafchamps, M. (2003) *Rural Poverty, Risk, and Development* Cheltenham, UK: Edward Elgar Publishing.

Fafchamps, M. (2010) "Vulnerability, Risk Management, and Agricultural Development", *African Journal of Agricultural Economics,* 5(1):243-260.

Haggblade, S., Hazell, P. and Reardon, T. (2007) *Transforming the Rural Nonfarm Economy.* Baltimore: Johns Hopkins University Press and International Food Policy Research Institute.

Investment: Theory and Evidence from Semi-Arid India," *Journal of Business and Economic Statistics*, 15(2): 180-194.

Jin, S. and Deininger, K. (2009) "Key Constraints for Rural Non-Farm Activity in Tanzania: Combining Investment Climate and Household Surveys" *Journal of African Economies* 18(2): 319-361

Kurosaki, T. and Fafchamps, M. (2002) "Insurance Market Efficiency and Crop Choices in Pakistan" *Journal of Development Economics* 67(2): 419-53.

Lanjouw, J. and Lanjouw, P. (2001) "The rural non-farm sector: issues and evidence from developing countries" *Agricultural Economics* 26(1): 1-23.

McKenzie, D. and Woodruff, C. (2006) "Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises" *Economic Development and Cultural Change* 55(1): 3-42.

Mead, D. C. and Liedholm, C. (1998) "The dynamics of micro and small enterprises in developing countries" *World Development* 26(1): 61-74.

Morduch, J. (1995) "Income Smoothing and Consumption Smoothing," *Journal of Economic Perspectives* 9(3): 103-14.

Morduch, J. (1999) "The Microfinance Promise," *Journal of Economic Literature* 37(4):1569-1614.

Patillo, C. (1998) "Investment, Uncertainty, and Irreversibility in Ghana" *IMF Staff papers* 45(3).

Paxson, P. and Schady, N. (2005) "Child Health and Economic Crisis in Peru," *World Bank Economic Review,* 19(2): 203-223.

Porter, C. (forthcoming). "Examining the impact of idiosyncratic and covariate shocks on Ethiopian households' consumption and income sources" *Journal of Development Studies.*

Rijkers, B. and Costa, R. (forthcoming) "Gender and Rural Non-Farm Entrepreneurship" *World Development.*

Rijkers, B., Söderbom, M., and Loening, J. (2010) "A Rural-Urban Comparison of Manufacturing Enterprise Performance in Ethiopia." *World Development* 38(9): 1278-1296.

Rose, E. (2001) "Ex ante and ex post labor supply response to risk in a low-income area" *Journal of Development Economics* 64 (2): 371–388.

Rosenzweig, R. and Binswanger, H. (1993) "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments" *The Economic Journal* 103(41):56-78.

Sen, A. (1981), *Poverty and Famines*, Clarendon Press, Oxford.

Söderbom, M. and Teal, F. (2004) "Size and Efficiency in African Manufacturing Firms: Evidence from Firm-Level Panel Data" *Journal of Development Economics* 73:369-394.

Townsend R.M. (1995) "Consumption Insurance: An Evaluation of Risk-bearing Systems in Low-Income Economies" *Journal of Economic Perspectives* 9:83-102.

Vargas Hill, R (2010) "Investment and Abandonment Behavior of Rural Households: An Empirical Investigation," *American Journal of Agricultural Economics* 92(4): 1065-1086

World Bank (2005) "Well-Being and Poverty in Ethiopia: The Role of Agriculture and Agency" Report No. 29468-ET. Poverty Reduction and Economic Management 2 (AFTP2), Country Department for Ethiopia, Africa Region.

Zwede and Associates (2002) *Jobs, Gender and Small Enterprises in Africa: Preliminary Report, Women Entrepreneurs in Ethiopia*. ILO Office, Addis Ababa in association with SEED, International Labour Office, Geneva, October.

[i] Shocks may force households to sell productive assets, withdraw children from school (see e.g. de Janvry et al., 2006), invest less in health and nutrition (see e.g. Paxson and Schady 2005, Dercon and Krishnan, 2000) or reduce essential consumption, all of which compromise future income earning potential by reducing households' physical and human capital stocks.

[ii] It should be noted, however, that Africa's growth performance in the recent past has been much stronger than during the 1980s and 1990s.

[iii] However, not all observations contained information on all variables of interest, which is why the numbers of observations used in our various analyses are lower than the number reported here.

[iv] Out of 9 product-measurement unit combinations only kakitala, locally produced beer, is correlated with wereda-level WRSI in 2006. However, it should be noted that the number of observations is limited in most cases and that the definitions of local units of measurement, such as a goug, differ across localities, which limits the reliability of these correlations.

[v] One potential concern with our measure of risk is that, if WRSI is trended, then risk may be better captured as the excess volatility in the de-trended series. We have examined whether an alternative measure of uncertainty, based on de-trended series, is very different from our original measure. Specifically, using the 1996-2005 sample, we first regressed actual WRSI on a full set of wereda dummies and interaction terms between time and the wereda dummies. We then computed the wereda specific standard deviation of the residual from that regression – which thus nets out wereda specific time trends - and compared that measure to our original measure. It turns out that the correlation between these two measures of risk is as high as 0.97. Thus, in our particular application, whether we net out wereda specific time trends or not before computing our risk measure matters little.

[vi] Variable definitions are provided in the Appendix.

[vii] 4% of all households own more than one enterprise. The maximum number of enterprises owned by a single household is 4 (we found 2 such households in the data).

[viii] We also examined the impact of risk and shocks on the number of enterprises owned by each household using ordered probit and Poisson regression models. The results, which we do not present to conserve space, are qualitatively very similar to the results obtained using the simple probit.

[ix] Similar robustness checks were run for the production function, investment and capital stock regressions presented below. In none of these models did we find evidence of bias associated with endogenous placement. Results are omitted to conserve space, but available from the authors upon request.

[x] The probability of running a firm is concavely related to the age of the household head. This may be picking up lifecycle effects or the fact that older individuals are usually wealthier and consequently better capable of raising the capital required to set up a profitable firm. The likelihood of participating is also concavely associated with the educational attainment of the household head, with the turning point at around 4 years of education. The fact that the probability of running an enterprise decreases as schooling increases beyond 4 years is driven by the fact that education provides better access to wage jobs, which tend to be better paid. The regressions also suggest that divorced individuals are much more likely to run a firm. Perhaps these individuals lack alternative income earning options, assets, and support from their family and friends.

[xi] Only 3 out of our 729 enterprises in the entire RICS-dataset employed more than 10 employees. However, these enterprises are all household-based and we might have missed out on fully commercial enterprises owned or managed by individuals not living in these communities. However, from the community level dataset one can infer that there are not more than a dozen firms with more than 20 employees in a radius of 1-hour commuting distance from the 179 surveyed communities. It thus seems safe to conclude that there are very few large firms in rural areas.

[xii] We also experimented with selection correction terms based on an ordered probit model for the number of non-farm enterprises. The results, which are not presented to conserve space but available from the authors upon request, were qualitatively similar to those obtained using the Heckman selection correction term.

[xiii] It is clearly possible that these exclusion restrictions are violated, for example if there is sorting by ability in the marriage market.

[xiv] We also experimented with the more flexible translog production function, which can be interpreted as a second-order Taylor approximation to a more general production function. Based on this framework, we did not reject the Cobb-Douglas restrictions for any of our estimations. We therefore proceed with the Cobb-Douglas framework, which is remarkably robust across African firm-level data. This facilitates interpretation of the results, and retains comparability with both micro- and macro-approaches that have been used to assess the impact of the investment climate (see e.g. Dollar et al., 2005) on firm performance.

[xv] We attempted to exploit information on prices to instrument factor choices, but the difficulties associated with constructing comparable price indicators were prohibitively formidable since units of measurement for different goods varied substantially across villages. Moreover, enterprises located in different villages produced different products, leaving us with a very small subsample of observations for which price variation could be identified.

[xvi] Of course, this does not imply that education necessarily boosts productivity. Education may be endogenous. For example, it could be the case that more able individuals are both more productive and more likely to set up a firm. Furthermore, the convex relationship between education and non-farm enterprise profitability may be driven by selection. For example, suppose that enterprise productivity is randomly distributed and that the likelihood of being offered a wage job increases with educational attainment. Suppose furthermore that wage opportunities are on average better paid than self-employment in non-farm enterprises. Most educated managers will take up wage jobs. However, those with particularly productive enterprises might prefer to reject wage job offers, since they do not yield a higher income stream. Such a selection process could thus cause a spurious convex correlation between education and profitability.

[xvii] The average of the log of investment for firms which invested is 4.77, corresponding to about 118 Birr, or about $13.

[xviii] We also ran regressions in which we used a selection correction based on an ordered probit model for the number of enterprises owned. The results, which are not presented to conserve space but available from the authors upon request, are qualitatively similar to those obtained using the Heckman procedure.

[xix] Thus our measures of the long-run capital stock and capital at startup are defined differently.

Table 1: Correlation Matrix – Predicted Agricultural Performance (WRSI 2006) and prices of individual products

		WRSI 2006 local	Tella (can)	Kakitala (goug)	Sugar (goug)	Tella (goug)	Coffee (kg)	Cereal (kg)	Soap	Scarfs
WRSI 2006	Correlation	**1**								
	p-value									
	Observations	672								
Tella (can)	Correlation	0.15	**1**							
	p-value	0.28								
	Observations	52	54							
Kakitala (goug)	Correlation	-0.32		**1**						
	p-value	0.02								
	Observations	49		63						
Sugar (goug)	Correlation	0.11			**1**					
	p-value	0.64								
	Observations	22			23					
Tella (goug)	Correlation	-0.10				**1**				
	p-value	0.56								
	Observations	33				37				
Coffee (kg)	Correlation	0.30			-0.12		**1**			
	p-value	0.27			0.67					
	Observations	16			15		16			
Cereal (kg)	Correlation	-0.07						**1**		
	p-value	0.81								
	Observations	13						14		
Soap	Correlation	-0.05			0.24		-0.10		**1**	
	p-value	0.79			0.30		0.75			
	Observations	27			19		12		29	
Scarfs	Correlation	0.01								**1**
	p-value	0.98								
	Observations	28								30

Note: Tella and kakitala are locally brewed alcoholic beverages. Goug is a local measure. However, the precise quantity of a goug varies across villages, which makes the price comparisons somewhat unreliable. Because of measurement issues, we have decided to consider each product-measurement unit combination as a separate product instead of attempting to convert different units into comparable quantities, which explains why a can of tella is not considered the same product as a goug of tella. Many rows are empty because not all products were produced in each wereda

Table 2: Household Characteristics by Enterprise Ownership

	All Households				Households owning an enterprise	Households without enterprise
	Mean	*SD*	*p5*	*p95*	*Mean*	*Mean*
Household Characteristics						
Age of the Head (Years)	43.41	16.14	22	74	42.87	43.60
Years of Schooling	1.84	3.54	0	9	1.72	1.88
Gender (1=male)	0.69				0.52	0.75
Married	0.66				**0.55**	**0.70**
Divorced or separated	0.14				**0.28**	**0.10**
Widowed	0.14				**0.14**	**0.14**
Migrant	0.34				**0.46**	**0.30**
Household size	4.19	2.20	1	8	**4.01**	**4.25**
Adult men (#)	1.27	1.07	0	3	**1.11**	**1.32**
Adult women (#)	1.37	0.86	0	3	**1.49**	**1.33**
Children (<5 years)	0.61	0.72	0	2	**0.51**	**0.64**
Elderly (>65 years)	0.15	0.39	0	1	0.13	0.16
Distance to the nearest market in km (log)	1.77	0.86	0.26	2.94	**1.37**	**1.91**
Distance to the nearest all-weather road in km	1.97	1.23	0.10	3.76	**1.54**	**2.12**
Community Characteristics						
Remote	0.57				**0.37**	**0.65**
Rural Town	0.22				**0.45**	**0.13**
Credit Institution	0.61				**0.72**	**0.57**
Land usable as collateral	0.13				0.15	0.12
Daily wage (log) (ETB)	2.19	0.33	1.61	2.77	**2.24**	**2.17**
Lack of electricity access	0.38	0.27	0.05	0.93	**0.34**	**0.39**
Uncertainty						
σ WRSI 1995-2005	7.43	6.09	1.00	14.41	7.57	7.38
Shocks and Coping Capacity						
WRSI 2006	98.28	2.65	92.75	100	98.27	98.28
Illness shock	0.20				**0.19**	**0.24**
Job-loss shock	0.02				0.01	0.02
Death shock	0.04				0.04	0.04
Emergency credit	0.59				**0.55**	**0.60**
Household Expenditure						
Log total household expenditure (ETB)	8.24	0.72	7.06	9.38	8.21	8.26
Log annual expenditure per adult (ETB)	7.31	0.63	6.30	8.33	7.33	7.30
Ratio total enterprise value added to total hh*	0.33	0.35	0.0025	1.00	0.33	
Participation Rate (weighted)		23%				
Observations		2009			523	1486

Note: Statistics are unweighted, with the exception of the participation rate. **Bolded** coefficients indicate differences between enterprise owning and non-enterprise owning households are significant at the 5% level. Exchange rate USD-Ethiopian Birr: 1:9.1. See the appendix for the definitions of these variables

* The statistics was computed using the sample of enterprise owning households only

Table 3. Participation Probits: Marginal Effects

	(1)	(2)	(3)	(4)
Uncertainty				
σ WRSI 1995-2005	0.003	0.001	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Shocks and coping capacity				
WRSI 2006	0.009**	0.007*	0.008*	0.008*
	(0.004)	(0.004)	(0.004)	(0.004)
Emergency credit			0.022	0.022
			(0.022)	(0.022)
Ilness shock			0.024	0.024
			(0.026)	(0.026)
Job loss shock			-0.028	-0.028
			(0.069)	(0.069)
Death shock			-0.045	-0.045
			(0.047)	(0.047)
Household Characteristics				
Head's Age	0.014***	0.015***	0.015***	0.015***
	(0.004)	(0.004)	(0.004)	(0.004)
Head's Age^2	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Head's Schooling	0.060***	0.060***	0.060***	0.060***
	(0.010)	(0.010)	(0.010)	(0.010)
Head's Schooling^2	-0.008***	-0.008***	-0.008***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Gender of the Head	-0.130***	-0.142***	-0.144***	-0.143***
	(0.041)	(0.042)	(0.042)	(0.042)
Married	0.076	0.071	0.067	0.067
	(0.053)	(0.053)	(0.054)	(0.054)
Divorced or separated	0.216***	0.197**	0.200**	0.200**
	(0.081)	(0.081)	(0.081)	(0.081)
Widowed	0.000	-0.017	-0.014	-0.014
	(0.067)	(0.065)	(0.066)	(0.066)
Migrant	0.044*	0.046**	0.046*	0.059*
	(0.023)	(0.023)	(0.024)	(0.035)
Share adult men	-0.012	-0.012	-0.013	-0.013
	(0.012)	(0.012)	(0.012)	(0.012)
Share adult women	0.038***	0.037***	0.036***	0.036***
	(0.013)	(0.013)	(0.013)	(0.013)
Share children (< 5 yrs)	-0.023	-0.023	-0.023	-0.023
	(0.015)	(0.015)	(0.015)	(0.015)
Share elderly (>65 yrs)	0.045	0.051	0.050	0.050
	(0.040)	(0.040)	(0.040)	(0.040)
Distance to food market (log)	-0.221***	-0.245***	-0.241***	-0.240***
	(0.053)	(0.054)	(0.054)	(0.054)
Distance to food market (log)^2	0.058***	0.063***	0.062***	0.062***
	(0.014)	(0.014)	(0.014)	(0.014)
Distance to nearest road (log)	-0.018	-0.027	-0.028	-0.028
	(0.037)	(0.038)	(0.038)	(0.038)
Distance to nearest road (log)^2	0.005	0.006	0.007	0.007
	(0.008)	(0.008)	(0.008)	(0.008)

The table continues on the next page.

Table 3 continued.

	(1)	(2)	(3)	(4)
Community Characteristics				
Remote	-0.055	-0.052	-0.053	-0.053
	(0.036)	(0.036)	(0.036)	(0.036)
Town	0.240***	0.207***	0.214***	0.215***
	(0.053)	(0.054)	(0.055)	(0.055)
Credit Institution		0.012	0.010	0.011
		(0.024)	(0.024)	(0.024)
Land usable as collateral		0.073**	0.075**	0.074**
		(0.036)	(0.036)	(0.036)
Lack of electricity access		-0.105***	-0.102***	-0.101***
		(0.038)	(0.038)	(0.038)
Daily wage (log)		-0.013	-0.018	-0.020
		(0.032)	(0.032)	(0.032)
Migrant* σ WRSI 1995-2005				-0.002
				(0.003)
Number of observations	2,009	2,009	2,009	2,009
Chi2 (df)	479 (21)	494 (25)	497 (29)	497 (30)
Pseudo R2	0.208	0.215	0.216	0.216

Note: The dependent variable is a dummy variable equal to one if the household runs a non-farm enterprise and zero otherwise. Standard errors, reported in parentheses, are heteroscedasticity robust and clustered at the community level. Significance at the 10%, 5% and 1% level is indicated by *, **, and ***, respectively.

Table 4. Enterprise Characteristics: Descriptive Statistics

	Mean	Sd	USD[3]
Firm age	8.88	9.90	
Sales (Y) & Value-Added (V)[1]			
Ln Y (ETB)	7.36	1.78	$173
Ln V) (ETB)	6.48	2.10	$72
Ln (V/L) (ETB/day)	1.22	1.83	$0.37
Factors of production[1]			
Ln M (ETB)	5.76	2.31	$35
Ln K (ETB)	5.10	2.41	$18
L (labor days per year)	297.38	398.67	
Ln(K/L) (ETB/day)	-0.15	2.13	
Competition[1]			
1-5 competitors	0.11	0.31	
> 5 competitors	0.27	0.44	
Investment[2]			
Any investment	0.18	0.39	
Log investment (if invested)	4.77	1.83	$13
Initial Conditions[2]			
Ln K – initial (ETB)	4.14	2.10	$6.90
Ln L – at startup (days)	5.06	1.11	
Size Distribution[2]			
One worker only (sole proprietorship)		74.3%	
Two workers		21.4%	
3 or 4 workers		3.2%	
5 and more		0.8%	
Sectoral Composition[2]			
Manufacturing		*56.1%*	
Manufacturing (excl: grain milling, food and beverages)		30.0%	
Food and beverages, brewing/distilling		24.8%	
Grain milling		1.2%	
Trade		*29.0%*	
Retail via stalls and markets		4.9%	
Retail (not stalls/mkts)		18.0%	
Services		*15.0%*	
Services		4.9%	
Transport services		1.2%	
Hotels and restaurants		8.4%	
Other (specialized services)		0.4%	

Note: See the appendix for variable definitions.

[1] Based on the sample on which the production function regressions presented in Table 5 are based, unweighted

[2] Based on the sample on which the investment probits presented in Table 6 are based, unweighted; note that these numbers include the owner amongst the workers.

[3] ETB-Dollar exchange rate: 9,1:1

Table 5. Estimated Production Functions

	(1) OLS	(2) Heckman	(3) Heckman	(4) Heckman
Uncertainty				
σ WRSI 1995-2005		0.001	0.001	-0.001
		(0.018)	(0.018)	(0.018)
Shocks				
WRSI 2006		0.098***	0.096***	0.096***
		(0.031)	(0.031)	(0.031)
Emergency credit				-0.136
				(0.150)
Ilness shock				0.053
				(0.163)
Job loss shock				-0.620
				(0.429)
Death shock				-0.197
				(0.383)
Factors of Production				
ln K	0.128***	0.118***	0.116***	0.116***
	(0.044)	(0.039)	(0.039)	(0.039)
ln L	0.654***	0.669***	0.693***	0.700***
	(0.120)	(0.096)	(0.096)	(0.096)
ln M	0.283***	0.295***	0.298***	0.300***
	(0.044)	(0.036)	(0.036)	(0.035)
Firm characteristics				
Food and beverages	0.152	0.111	0.090	0.132
	(0.239)	(0.216)	(0.216)	(0.217)
Grain milling	-1.104	-1.254*	-1.187*	-1.103
	(1.027)	(0.723)	(0.717)	(0.719)
Whole sale trade	1.067***	1.197***	1.084***	1.083***
	(0.263)	(0.384)	(0.388)	(0.387)
Retail via stalls and markets	0.850***	0.761*	0.742*	0.776*
	(0.299)	(0.428)	(0.424)	(0.423)
Retail (not stalls/mkts)	0.643**	0.682**	0.658**	0.696**
	(0.257)	(0.268)	(0.271)	(0.272)
Services	0.303	0.252	0.316	0.309
	(0.312)	(0.345)	(0.348)	(0.346)
Hotels and restaurants	0.353	0.336	0.307	0.340
	(0.293)	(0.270)	(0.270)	(0.270)
Transport services	1.226*	1.163	1.387*	1.455*
	(0.660)	(0.777)	(0.786)	(0.786)
Others(specialized services)	0.854***	0.919	1.130	1.235
	(0.246)	(1.292)	(1.292)	(1.290)
Firm age	0.010	0.009	0.008	0.008
	(0.006)	(0.007)	(0.007)	(0.007)
Distance to food market (log)	-0.012	0.231	0.288	0.267
	(0.289)	(0.386)	(0.381)	(0.382)
Distance to food market (log)^2	-0.006	-0.052	-0.051	-0.052
	(0.110)	(0.108)	(0.107)	(0.107)
Distance to nearest road (log)	0.378	0.402	0.370	0.365
	(0.254)	(0.257)	(0.254)	(0.253)
Distance to nearest road (log)^2	-0.066	-0.084	-0.069	-0.065
	(0.062)	(0.063)	(0.063)	(0.063)

The table continues on the next page.

Table 5 continued.

	(1)	(2)	(3)	(4)
Manager Characteristics				
Manager's age	-0.031	-0.034	-0.037	-0.036
	(0.024)	(0.025)	(0.025)	(0.025)
Manager's age2/100	0.212	0.233	0.268	0.251
	(0.222)	(0.256)	(0.256)	(0.256)
Manager's gender	0.565***	0.672***	0.605***	0.637***
	(0.187)	(0.190)	(0.189)	(0.192)
Manager's schooling	-0.191**	-0.244***	-0.231**	-0.215**
	(0.094)	(0.093)	(0.092)	(0.092)
Manager's schooling2/100	21.513*	27.809**	26.750**	24.557**
	(12.999)	(11.748)	(11.651)	(11.710)
Migrant	-0.163	-0.244	-0.218	-0.219
	(0.144)	(0.148)	(0.148)	(0.148)
Community Characteristics				
Remote	0.293	0.269	0.070	0.106
	(0.365)	(0.329)	(0.334)	(0.336)
Rural town	0.963***	0.732**	0.622**	0.652**
	(0.237)	(0.302)	(0.305)	(0.305)
Lack of electricity access			-0.337	-0.362
			(0.297)	(0.296)
Daily wage (log)			0.422*	0.440*
			(0.238)	(0.237)
1-5 competitors			-0.277	-0.253
			(0.227)	(0.230)
>5 competitors			0.142	0.148
			(0.166)	(0.166)
Constant	1.158	-7.994**	-8.814**	-8.836***
	(0.906)	(3.285)	(3.431)	(3.429)
Inverse Mills Ratio		-0.483*	-0.395	-0.355
		(0.261)	(0.262)	(0.263)
R2	0.469			
Adjusted R2	0.431			
N	374	1,877	1,877	1,877
Chi2 (df)		347(27)	364(31)	370 (35)

Note: The dependent variable is the logarithm of sales. The omitted competition category is "no competition"; the omitted sector is "Manufacturing". Standard errors, reported in parentheses, are heteroscedasticity robust and clustered at the community level. Significance at the 10%, 5% and 1% level is indicated by *, **, and ***, respectively. The Heckman model results were obtained by means of Heckman's two-step procedure, using the specification presented in table 3 column 2 as the participation equation.

Table 6. Selection Corrected Investment Probits: Parameter Estimates

	(1)	(2)	(3)	(4)
Uncertainty				
σ WRSI 1995-2005	-0.052***	-0.053***	-0.054***	-0.056***
	(0.015)	(0.014)	(0.015)	(0.016)
Shocks				
WRSI 2006	-0.034	-0.037	-0.038	-0.050
	(0.034)	(0.034)	(0.034)	(0.034)
Emergency credit			0.590***	0.561***
			(0.165)	(0.167)
Ilness shock			0.138	0.181
			(0.166)	(0.168)
Job loss shock			0.301	0.351
			(0.412)	(0.415)
Death shock			0.026	0.024
			(0.360)	(0.358)
Firm Characteristics				
Food and beverages	0.102	0.023	-0.069	-0.248
	(0.197)	(0.196)	(0.197)	(0.206)
Grain milling	1.111**	1.073**	0.814	0.108
	(0.514)	(0.516)	(0.498)	(0.621)
Whole sale trade	-0.331	-0.500	-0.535*	-1.016***
	(0.320)	(0.323)	(0.311)	(0.339)
Retail via stalls and markets	0.246	0.168	0.194	0.108
	(0.302)	(0.302)	(0.310)	(0.340)
Retail (not stalls/mkts)	0.060	-0.019	-0.148	-0.509**
	(0.217)	(0.216)	(0.220)	(0.256)
Services	-0.176	-0.301	-0.299	-0.296
	(0.296)	(0.300)	(0.287)	(0.302)
Hotels and restaurants	0.567**	0.546**	0.433	0.259
	(0.268)	(0.262)	(0.267)	(0.272)
Transport services	0.479	0.345	0.136	-0.511
	(0.506)	(0.520)	(0.517)	(0.540)
Firm age	0.016**	0.014*	0.012	0.014*
	(0.007)	(0.007)	(0.008)	(0.008)
Distance to food market (log)	-0.490	-0.300	-0.147	-0.141
	(0.411)	(0.547)	(0.493)	(0.516)
Distance to food market (log)^2	0.194*	0.171	0.144	0.142
	(0.111)	(0.140)	(0.130)	(0.136)
Distance to nearest road (log)	-0.061	-0.085	0.025	0.142
	(0.244)	(0.260)	(0.251)	(0.254)
Distance to nearest road (log)^2	-0.004	-0.003	-0.033	-0.062
	(0.063)	(0.063)	(0.062)	(0.063)

The table continues on the next page.

Table 6 continued.

	(1)	(2)	(3)	(4)
Manager characteristics				
Manager's age	0.029	0.016	0.013	0.008
	(0.030)	(0.032)	(0.032)	(0.032)
Manager's age2/100	-0.341	-0.199	-0.166	-0.103
	(0.299)	(0.328)	(0.328)	(0.325)
Manager's gender	0.369*	0.434*	0.275	0.145
	(0.216)	(0.230)	(0.228)	(0.248)
Manager's schooling	0.038	0.005	-0.027	-0.055
	(0.098)	(0.110)	(0.100)	(0.105)
Manager's schooling2/100	-0.082	3.566	8.011	11.586
	(12.023)	(13.001)	(11.638)	(12.181)
Migrant	0.116	0.111	0.099	0.061
	(0.146)	(0.159)	(0.156)	(0.162)
Community Characteristics				
Remote	-0.148	-0.167	-0.259	-0.173
	(0.326)	(0.334)	(0.332)	(0.340)
Rural town	0.050	0.134	0.138	0.078
	(0.306)	(0.317)	(0.324)	(0.322)
Lack of electricity access		-0.145	-0.065	-0.205
		(0.289)	(0.290)	(0.302)
Daily wage (log)		0.045	0.009	0.020
		(0.232)	(0.232)	(0.236)
1-5 competitors		-0.018	-0.011	-0.044
		(0.221)	(0.228)	(0.230)
>5 competitors		0.303*	0.257	0.163
		(0.160)	(0.157)	(0.157)
Credit Institution		-0.438**	-0.430**	-0.424**
		(0.199)	(0.200)	(0.204)
Land usable as collateral		-0.098	-0.075	-0.170
		(0.253)	(0.242)	(0.253)
Initial Conditions				
Ln K – initial				0.184***
				(0.048)
ln L – at startup				0.089
				(0.072)
Constant	1.982	2.797	2.679	2.978
	(3.766)	(3.743)	(3.735)	(3.795)
Inverse Mills Ratio	-0.090	-0.263	-0.330	-0.353
	(0.373)	(0.429)	(0.339)	(0.360)
N	2,085	2,085	2,085	2,085
chi2 (df)	50(23)	67(29)	81(33)	102(35)

Note: The dependent variable is a dummy variable equal to one if the enterprise has made any investment and zero otherwise. The omitted competition category is "no competition"; the omitted sector is "Manufacturing". The category "Other" dropped out because of multicollinearity. Standard errors, reported in parentheses, are heteroscedasticity robust and clustered at the community level. Significance at the 10%, 5% and 1% level is indicated by *, **, and ***, respectively. The results were obtained by means of Heckman's two-step procedure, using the specification presented in table 3 column 2 as the participation equation.

Table 7. The determinants of the capital stock

	(1)	(2)	(3)	(4)	(5)	(6)
Uncertainty						
σ WRSI 1995-2005	-0.051*	-0.058**	-0.058**	-0.040	-0.027	-0.021
	(0.028)	(0.028)	(0.027)	(0.027)	(0.024)	(0.026)
Shocks						
WRSI 2006	0.032	0.022	0.017	0.021	0.033	0.023
	(0.053)	(0.053)	(0.053)	(0.047)	(0.045)	(0.042)
Emergency credit			0.956***	0.624***	0.486**	0.365*
			(0.245)	(0.227)	(0.210)	(0.199)
Ilness shock			-0.141	-0.053	-0.103	-0.072
			(0.296)	(0.247)	(0.238)	(0.209)
Job loss shock			-0.539	-0.296	-0.538	-0.276
			(0.628)	(0.558)	(0.396)	(0.390)
Death shock			1.156	1.186	0.581	0.688
			(0.990)	(0.731)	(0.630)	(0.501)
Firm Characteristics						
Grain milling					5.526***	3.781***
					(0.797)	(0.640)
Whole sale trade					1.071*	0.487
					(0.560)	(0.662)
Retail via stalls and markets					-0.760	-0.878*
					(0.473)	(0.524)
Retail (not stalls/mkts)					0.565	-0.169
					(0.360)	(0.358)
Services					-0.531	-0.615
					(0.574)	(0.468)
Hotels and restaurants					2.592***	2.160***
					(0.328)	(0.306)
Transport services					3.585***	1.882***
					(0.404)	(0.448)
Food and beverages					2.338***	1.944***
					(0.283)	(0.258)
Firm age	0.007	0.005	0.002	0.017*	0.019*	0.023***
	(0.011)	(0.011)	(0.011)	(0.009)	(0.010)	(0.009)
Distance to food market	-1.670**	-0.645	-0.522	-1.089*	-0.953*	-1.117**
(log)	(0.672)	(0.901)	(0.760)	(0.577)	(0.506)	(0.451)
Distance to food market	0.337*	0.148	0.178	0.303**	0.343**	0.369***
$(\log)^2$	(0.191)	(0.223)	(0.198)	(0.154)	(0.142)	(0.124)
Distance to nearest road	0.068	0.163	0.266	0.317	-0.323	-0.132
(log)	(0.431)	(0.473)	(0.461)	(0.378)	(0.364)	(0.325)
Distance to nearest road	-0.061	-0.090	-0.121	-0.140	0.015	-0.037
$(\log)^2$	(0.103)	(0.111)	(0.109)	(0.089)	(0.088)	(0.078)

The table continues on the next page.

Table 7 continued.

	(1)	(2)	(3)	(4)	(5)	(6)
Manager Characteristics						
Manager's age	0.002	-0.024	-0.038	-0.038	-0.020	-0.028
	(0.041)	(0.046)	(0.041)	(0.036)	(0.034)	(0.031)
Manager's age2/100	-0.212	0.056	0.237	0.289	0.149	0.259
	(0.412)	(0.470)	(0.417)	(0.356)	(0.340)	(0.315)
Manager's gender	0.164	0.429	0.181	-0.238	0.890***	0.583**
	(0.311)	(0.377)	(0.334)	(0.275)	(0.245)	(0.230)
Manager's schooling	0.375**	0.229	0.156	0.093	0.057	0.027
	(0.153)	(0.184)	(0.162)	(0.136)	(0.125)	(0.120)
Manager's schooling2/100	-32.615*	-17.056	-8.174	-5.588	-2.999	-1.226
	(19.072)	(22.910)	(20.525)	(17.983)	(15.827)	(15.637)
Migrant	0.165	0.070	0.055	0.022	0.207	0.152
	(0.255)	(0.283)	(0.259)	(0.226)	(0.198)	(0.187)
Community Characteristics						
Remote	0.325	0.128	-0.217	0.165	-0.645	-0.428
	(0.537)	(0.561)	(0.563)	(0.504)	(0.477)	(0.458)
Rural town	-0.226	-0.107	-0.061	-0.418	-0.329	-0.577
	(0.545)	(0.538)	(0.538)	(0.474)	(0.459)	(0.406)
Lack of electricity access		-0.837*	-0.655	-0.687	-0.122	-0.300
		(0.505)	(0.511)	(0.445)	(0.436)	(0.408)
Daily wage (log)		0.056	0.044	0.270	0.245	0.322
		(0.386)	(0.365)	(0.312)	(0.304)	(0.284)
1-5 competitors		-0.007	-0.135	-0.268	0.117	0.039
		(0.439)	(0.407)	(0.341)	(0.359)	(0.327)
>5 competitors		0.259	0.211	0.008	0.269	0.170
		(0.251)	(0.244)	(0.220)	(0.203)	(0.189)
Credit Institution		-0.709**	-0.708*	-0.619*	-0.722**	-0.667**
		(0.358)	(0.369)	(0.318)	(0.312)	(0.287)
Land usable as collateral		-0.579	-0.518	-0.623*	-0.170	-0.290
		(0.515)	(0.448)	(0.361)	(0.302)	(0.277)
Initial Conditions						
Ln K – initial				0.589***		0.456***
				(0.068)		(0.066)
ln L – at startup				0.178**		0.095
				(0.078)		(0.076)
Constant	3.335	5.468	5.834	1.988	1.572	0.843
	(5.447)	(5.544)	(5.412)	(4.759)	(4.646)	(4.328)
Inverse Mills Ratio	0.209	-0.449	0.569	-0.419	-0.123	-0.128
	(0.539)	(0.771)	(0.595)	(0.381)	(0.276)	(0.222)
N	1,909	1,909	1,909	1,909	1,909	1,909
chi2 (df)	51 (15)	71 (21)	94 (25)	210 (27)	431(33)	500(35)

Note: The dependent variable is the logarithm of the capital stock. The omitted competition category is "no competition"; the omitted sector is "Manufacturing". The category "Other" dropped out because of multicollinearity. Standard errors, reported in parentheses, are heteroscedasticity robust and clustered at the community level. Significance at the 10%, 5% and 1% level is indicated by *, **, and ***, respectively. The results were obtained by means of Heckman's two-step procedure, using the specification presented in table 3 column 2 as the participation equation.

Table 8. The determinants of the capital-labour ratio

	(1)	(2)	(3)	(4)	(5)	(6)
Uncertainty						
σ WRSI 1995-2005	-0.052*	-0.061**	-0.062**	-0.044	-0.031	-0.025
	(0.030)	(0.029)	(0.028)	(0.028)	(0.026)	(0.027)
Shocks						
WRSI 2006	0.027	0.012	0.007	0.003	0.024	0.007
	(0.052)	(0.052)	(0.051)	(0.047)	(0.044)	(0.042)
Emergency Credit			0.781***	0.578**	0.306	0.307
			(0.242)	(0.226)	(0.211)	(0.199)
Job loss shock			-0.218	-0.170	-0.172	-0.187
			(0.283)	(0.245)	(0.235)	(0.208)
Death shock			-0.507	-0.209	-0.557	-0.208
			(0.750)	(0.585)	(0.528)	(0.414)
Illness shock			0.854	1.271*	0.334	0.760
			(0.934)	(0.724)	(0.612)	(0.505)
Inverse Mills Ratio	0.127	-0.482	-0.561	-0.243	-0.189	-0.156
	0.477	0.527	0.451	0.208	0.267	0.230
Basic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Additional IC controls	No	Yes	Yes	Yes	Yes	Yes
Shocks	No	No	Yes	Yes	Yes	Yes
Initial Capital	No	No	No	Yes	No	Yes
Subsector Dummies	No	No	No	Yes	Yes	Yes
N	1,909	1,909	1,909	1,909	1,909	1,909
chi2 (df)	46(15)	65(21)	79(25)	564 (25)	229 (27)	564 (24)

Note: The dependent variable is the logarithm of the capital-labour ratio. The omitted competition category is "no competition"; the omitted sector is "Manufacturing". The category "Other" dropped out because of multicollinearity. Standard errors, reported in parentheses, are heteroscedasticity robust and clustered at the community level. Significance at the 10%, 5% and 1% level is indicated by *, **, and ***, respectively. The results were obtained by means of Heckman's two-step procedure, using the specification presented in table 3 column 2 as the participation equation.

Table 9. The determinants of the capital stock: Results from Quantile Regressions

	(1) 10th percentile	(2) 30th percentile	(3) Median	(4) 70th percentile	(5) 90th percentile
Uncertainty					
σ WRSI 1995-2005	-0.026	-0.020	-0.060	-0.079**	-0.073**
	(0.046)	(0.044)	(0.037)	(0.036)	(0.036)
Shocks					
WRSI 2006	-0.032	0.039	0.031	0.047	-0.003
	(0.079)	(0.076)	(0.063)	(0.061)	(0.061)
Basic Controls	Yes	Yes	Yes	Yes	Yes
N	413	413	413	413	413

Note: The specifications replicate that of column 1 in table 7 with the selection term excluded. Significance at the 10%, 5% and 1% level is indicated by *, **, and ***, respectively. The controls that are included but not presented are: Firm age, distance to food market (log) and its square, distance to nearest road and its square, manager's age and its square, manager's gender, manager's schooling, as well as dummies for remote locations and rural towns.

Appendix: Variable Definitions

Credit Institution: dummy variable that takes the value one if a credit institution is present in the village (and zero otherwise).

Daily wage (log): daily wage of a male agricultural casual labourer – extracted from the community survey.

Emergency credit: dummy variable that takes the value 1 if the household indicated it was able to access 100 Birr in case of an emergency, and zero otherwise.

Lack of electricity access: proportion of firms (excluding firm i) in the community who indicate that they are not using electricity because it is unavailable.

Land usable as collateral: dummy variable that takes the value 1 if there is a credit institution in the village and land can be used as collateral.

Ln K: the log of the replacement value of all machinery and equipment used by the firm.

Ln K – initial: initial capital, measured as the answer to the question "how much money did you pay to set up the firm?"

Ln L: log total labor usage measured in labor days.

Ln L - startup: log total labor usage at startup measured in labor days.

Ln M: the log of the sum of the value of all material inputs used and the value of all products purchased for resale.

Ln Y: log total enterprise sales over the year preceding the survey.

Migrant: dummy value that takes the value 1 if the household head was not born in the community he/she is currently living in and zero otherwise.

Remote: a dummy that takes the value 1 if the household/enterprise located in a remote rural area, following the Andersson and Jirström (2007) discrete urbanization index.

Rural town: a dummy that takes the value 1 if the household/enterprise is located in a rural town.

Shock (Job-loss /Death /illness): dummy variable taking the value 1 if one of the household members suffered a shock (i.e. suffered a job loss/died /suffered a severe illness) in the year preceding the survey.

WRSI 2006: The Water Requirement Satisfaction Index (WRSI) in 2006 is an indicator of predicted crop performance based on water availability during the growing season, computed using rainfall information and measured at the wereda (i.e. district) level. WRSI is expressed as a percentage of maximum yield under ideal growing conditions. WRSI data were obtained from National Oceanic and Atmospheric Administration's Africa Rainfall Estimates Climatology.

σ WRSI 1995-2005: Standard deviation of wereda WRSI (see above) over the period 1995-2005.

1-5 competitors; dummy variable equal to 1 if the firm faces 1-5 competitors in the community in which it operates.

>5 competitors: dummy variable equal to 1 if the firm faces more than 5 competitors in the community in which it operates.