Constrained firms, not subsistence activities: Evidence on capital returns and accumulation in Peruvian microenterprises

Marina Dodlova
GIGA German Institute of Global and Area Studies, Hamburg, Germany
CESifo

Kristin Göbel
University of Hamburg, Germany

Michael Grimm
University of Passau, Germany
Erasmus University Rotterdam, The Netherlands

Jann Lay (corresponding author)
University of Göttingen, Germany
GIGA German Institute of Global and Area Studies, Hamburg, Germany

Neuer Jungfernstieg 21 / 20354 Hamburg / Germany
Tel.: ++49-(0)40-42825-763
Fax: ++49-(0)40 - 428 25-547
E-Mail: jann.lay@giga-hamburg.de

Acknowledgements: This research is part of a project entitled "Unlocking potential: Tackling economic, institutional and social constraints of informal entrepreneurship in Sub-Saharan Africa" (http://www.iss.nl/informality) funded by the Austrian, German, Norwegian, Korean and Swiss Government through the World Bank's Multi Donor Trust Fund Project: "Labor Markets, Job Creation, and Economic Growth, Scaling up Research, Capacity Building, and Action on the Ground". The financial support is gratefully acknowledged. The project is led by the International Institute of Social Studies of Erasmus University Rotterdam, The Hague, The Netherlands. The other members of the research consortium are: AFRISTAT, Bamako, Mali, DIAL-IRD, Paris, France, the German Institute of Global and Area Studies, Hamburg, Germany and the Kiel Institute for the World Economy, Kiel, Germany. We thank participants at the 6th IZA/World Bank Conference: Employment and Development in Mexico City, the Development Economics Conference 2011 of the German Economic Association in Berlin, and the UNU-WIDER Nordic Conference in Development Economics 2011 in Copenhagen for useful comments and suggestions.

Abstract

We present a multi-period model of capital accumulation in order to investigate microenterprise (ME) production dynamics in a developing country context characterized by credit constraints and risk. These constraints are reflected in marginal returns to capital above market interest rates and we show that capital accumulation is faster in MEs with higher productivity, higher initial wealth and less risk exposure. We test our predictions using panel data for Peruvian MEs from 2002 to 2006. We indeed find high marginal returns to capital and sizable effects of household non-business wealth and, in contrast to previous studies, risk on capital accumulation in MEs. The findings of this paper can serve as a basis for the promotion of combined credit and risk management devices to enhance private sector development.

Highlights

- We build a model of microenterprise (ME) dynamics under credit constraints and risk.
- Using panel data for Peruvian MEs we show high marginal returns to capital.
- Higher productivity and wealth, and less risk exposure foster capital accumulation.
- Effects of all constraints, including risk, are sizable, and there is an interaction between risk and wealth.
- This is evidence of constrained MEs rather than subsistence activities.

Key words: Microenterprises, capital accumulation, credit constraints, risk, firm growth, Peru.

JEL Codes:D13, D61, O12.

1. Introduction

Income from microenterprises (MEs) constitutes the main source of income of the growing number of urban dwellers in developing countries. It is widely assumed that some of these MEs may not realize their full earnings potential due to limited access to credit. This assumption provides the rationale for a key policy intervention towards MEs, micro-credit programs. Several empirical studies have found very high marginal returns to capital in MEs in poor countries – typically well above market interest rates, and in some studies highest at very low levels of capital stock (Fafchamps et al., 2011; Grimm, Krüger, and Lay, 2011; Kremer, Lee, and Robinson, 2010; McKenzie and Woodruff, 2006). This finding hints indeed at credit constraints as one major obstacle to firm growth. Otherwise, a profit-maximizing entrepreneur would increase the capital stock until marginal returns equalize the market interest rate. Empirical work that explicitly addresses credit constraints also confirms their importance (De Mel, McKenzie, and Woodruff, 2008; De Mel, McKenzie, and Woodruff, 2012; McKenzie and Woodruff, 2008; Dromel, Kolakez and Lehmann, 2010). De Mel et al. (2008), for instance, perform an experiment, in which they randomly provide cash or in-kind transfers to Sri Lankan MEs. The authors find, at least for male entrepreneurs, marginal returns to capital in a range from 55 to 70 percent per year. Consistent with credit market constraints, the marginal returns to capital are lower (higher) for wealthier (poorer) entrepreneurs. In a similar experiment in Mexico, McKenzie and Woodruff (2008) also provide evidence of very high returns of at least 20–33 percent a month. Returns are – with up to 100 percent monthly – highest among firms with capital stocks near $200 that report being financially constrained. At least for this selected group, there is hence a strong indication for credit market imperfections causing high returns at low levels of capital.[1]

Yet, credit constraints may only partly explain the observed high returns to capital. Successful entrepreneurs should be able and willing to re-invest a part of their high returns into their ME, thereby overcoming credit constraints, accumulating capital and bringing down marginal returns (as shown, for example, by Tybout, 1983). Risk may be the reason why this is not the case, as optimal capital stocks may be much lower in risky environments. The empirical literature generally has difficulties in the operationalization of risk and risk attitudes, which may explain why so far empirical research on the role of risk for ME performance is scarce. An exception is De Mel et al. (2008) who do, however, not find any sizeable effect of risk on returns to capital and no significant differences in marginal returns between risk-averse and less risk-averse entrepreneurs. Nevertheless, there is indirect evidence of the role of risk, such as the high rates of churning among informal MEs in developing countries (Mead and Liedholm, 1998).

In this paper we consider a multi-period model of capital accumulation that helps to disentangle the role of credit and risk constraints for ME production dynamics. In particular, we first consider the marginal returns to capital in a dynamic perspective. Second, by specifying utility and production functions we study how credit constrains and risky

[1] Similar evidence (for selected groups of entrepreneurs) is provided by Fafchamps et al. (2011).

environments separately as well as their interaction affect capital accumulation and may hamper ME production growth. Third, we examine the role of enterprise age in partly overcoming the effects of risk and credit constraints. Risk may be reduced through learning over time and credit constraints may be relaxed through savings.

Empirically, we test our model on the basis of a unique panel dataset of Peruvian MEs covering the period 2002 to 2006. The analysis of the ME sector is particularly relevant in the case of Peru, as these firms account for about 75 percent of urban employment in this country. Our data include mostly self-employed workers and firms with unpaid employees. Only 12% of all firms have paid employees. This is consistent with the theoretical framework where we consider a household production model. Even during the impressive growth period covered by our panel, with growth averaging 5.7 percent per annum, there has been no decline in employment in MEs. We first estimate production functions to test whether Peruvian MEs have high returns to invested capital. In addition, we use this step to compute total factor productivities and an idiosyncratic risk proxy. We then examine the determinants of capital accumulation by focusing on the two fundamental constraints to capital growth in MEs, credit constraints and risk. First, we confirm our model's prediction that marginal returns to capital for Peruvian MEs are higher due to credit and risk constraints. Second, both theoretically and empirically we find that the start-up capital is significantly lower and capital is accumulated more slowly due to credit constraints and risk. Finally, we show that credit constraints and risk interact and that the effect of risk and not credit constraints is less evinced in older enterprises.

This paper hence adds new insights on the returns to capital and the dynamics of small-scale production activities in developing countries. In addition to confirming high marginal returns, we enrich these findings on the particular and mixed effects of credit constraints and risky environments for ME performance. More importantly, our paper is one of rare papers presenting a dynamic theoretical analysis of capital accumulation for MEs. A unique panel data set on Peruvian MEs allows for both structurally testing our theoretical model predictions and providing primary evidence on capital accumulation. Our main contribution is hence to complement the static investigation of high returns to capital by a dynamic analysis of the determinants of capital accumulation with a focus on risk as well as to examine the effect of the interaction between credit and risk constraints.

The remainder of the paper is organized as follows. The subsequent section reviews previous findings in the literature. Section three presents a multi-period model of capital accumulation for MEs and derives the theoretical hypotheses. Section four describes the dataset and the basic characteristics of Peruvian micro-entrepreneurs and their enterprises. Section five presents our empirical setup and discusses the results. The final section summarizes our main findings and concludes.

2. Related literature

Our analysis adds to a growing recent literature on the performance of MEs in developing countries. Beyond the empirical studies that focus on returns to capital and their causes mentioned above, there is a separate strand of the literature that examines the patterns of entry, exit, and growth of MEs. Mead and Liedholm (1998) summarize the findings of a research project on ME behavior which draws (partly) on panel datasets from a number of developing countries. In most cases the authors find high rates of churning among MEs, with survival being positively associated with firm age, smaller initial size and past growth. The analysis of firm growth shows that MEs, which were smaller at start-up, tend to grow more rapidly than their larger counterparts. Moreover, young firms grow faster.[2] These results are similar to those obtained by Fajnzilber, Maloney, and Montes-Rojas (2006) using Mexican data. They find that size and time in business are negatively related to exit and growth. They conclude that MEs in Mexico show dynamic patterns consistent with a number of standard results from the theoretical literature on firm dynamics. Using a database of all registered firms in Cote d'Ivoire from 1977 to 1997 Klapper and Richmond (2011) confirm that the probability of survival increases with firm size, however, this effect is eliminated after 1987, the year which marks the start of the Ivorian Crisis.

While Mead and Liedholm (1998) and Fajnzilber et al. (2006) describe quite well some dynamic features of MEs in developing countries, they do not address the causes of differences in behavior.[3] Few studies explicitly show the impact of capital constraints on accumulation. One exception is a study about the garment industry in Tirupur in Southern India by Banerjee and Munshi (2004). The authors find large and systematic differences in the level of capital stock in firms owned by people from two different community groups. This finding may be attributed to differences in access to capital between the groups. One of them, the Gounders, comes from a relatively wealthy agricultural community that was the first to move into the garment industry in Tirupur. Banerjee and Munshi (2004) argue that the incumbent Gounders start their businesses with much higher levels of capital stock than comparable outsiders because of their stronger ties to the local community, and the associated better access to finance. Both groups accumulate capital over time, but the outsiders do so much faster to catch up with the Gounders after approximately seven years.[4]

---

[2] McPherson (1996) and Coad and Tamvada (2012), for example, show similar patterns for southern African and Indian firms, respectively.

[3] Fajnzylber et al. (2006) do present some suggestive evidence in favor of credit constraints. The authors regress employment growth on dummies for credit at start-up, and dummies for subsequent credit (and a set of other controls). Firms with start-up credit appear to grow slower given that they reach their optimal capital stocks more rapidly. In contrast, MEs with subsequent access to credit grow faster, as they can quickly adjust to their optimal capital stock. Yet, the authors acknowledge that this result might be driven by simultaneity.

[4] Further evidence for credit constraints is derived from exogenous (liquidity or credit availability) shocks on firm performance. Tybout (1983), for example, uses industry level data from Colombia. While favorable earnings shocks appear to be little relevant for large firms, these additional internal funds are used for investment by smaller firms. Banerjee and Duflo (2014) take advantage of a policy change in India that affected

Recent work by de Mel et al. (2012) – again based on the sample of Sri Lankan MEs – examines the long-term effects of the randomly assigned transfer of cash or capital. Tracking MEs over a period of 4.5 to 5.5 years they show that these transfers lead to permanently higher capital stocks and, accordingly, higher profits in treated MEs. De Mel et al. (2012) interpret these findings as an indication of tight credit constraints. Yet, the lack of re-investment of profits by the control group suggests that other factors must be at work as well. De Mel et al. (2012) point at behavioral factors, in particular the lack of self-control and time-inconsistent preferences, as a possible explanation why MEs do not accumulate capital – despite possibly high returns to investment.

Another explanation for the lack of (re-) investment is that profits are high on average, but highly risky and indeed MEs in developing countries take their economic decisions under high uncertainty and risks. Theses business risks come from a variety of sources and can be market, personal, natural, technological, and strategic ones (see Dunn et al., 1996). All these risks can be divided into two types that concern either price or output. While market risks refer to price uncertainty, personal risks, for example illness or death of the entrepreneur or other household members, can have important negative impacts on production, in particular in the context of MEs Natural risks including floods, droughts and other natural shocks also directly affect output. Similarly, technological risks can be associated with output losses, for example due to failures of new equipment. Strategic risks are caused by information asymmetries and opportunistic behavior by business partners and can, in principle, influence both price and output. While the greater portion of price risk is common to all firms – at least in a given timeframe and in a given sector – production risks have a more important idiosyncratic component. As we discuss in the empirical part below, it is challenging to empirically disentangle idiosyncratic and common business risks. In our theoretical framework below we model risk as production risk without distinguishing idiosyncratic or common risk factors.

Finally, one important caveat to the above finding of high marginal returns to capital stocks is that they tend to decline very fast with higher capital stocks (McKenzie and Woodruff, 2008; Grimm et al., 2011). This implies that – while capital scarcity may cause high marginal returns initially – the lack of (further) capital accumulation in MEs may be due to low overall productivity.

3. Theoretical framework

In this section, we briefly illustrate how credit constraints and risk explain the high returns to capital, as well as their possible impacts on capital accumulation. We present a multi-period model of capital stock accumulation, based on the elements of a static model by De Mel et al. (2008), and derive the hypotheses, which will be tested subsequently.

---

the flow of directed credit to estimate the effect of a favorable credit shock on investment and productivity of medium-sized firms. Their results show a large acceleration in the rate of growth of sales and profits due to the shock. In contrast, Akoten, Sawada, and Otsuka (2006) cannot find a positive impact of credit access on the performance of small garment producers in Kenya.

Without credit and risk constraints

Consider a multi-period model of the microenterprise, in which labor supply is exogenous and inelastically provided. Suppose that the microenterprise lives for an infinite number of periods. The household's endowment equals A. The number of other working adults in the household is denoted by n, and they are paid a fixed wage w. The household can make investments in its production through the formal credit market by borrowing (B), and through its internal household capital market, by allocating a part of the initial endowment (A_K) and the household's labor income to increase the capital stock. We assume that the household invests all labor income produced in period t in physical capital. The first-period capital stock equals K_0. Then, the household maximizes the intertemporally additive expected utility function:

$$E_0\left[\sum_{t=0}^{\infty}\beta^t u(C_t)\right] \to \max_{\{C_t\}_{t=0}^{\infty}}$$

where $0 < \beta < 1$ is a subjective intertemporal discount factor and C_t is consumption in period t subject to the capital accumulation equation

$$C_t = Y_t - K_{t+1} + r^t\left(A - A_{K_0}\right) - \sum_{i=1}^{t-1} r^{t-i} A_{K_i} \qquad (1)$$

where r is the interest rate for savings, the same for any time period, and $Y_t = f(K_t, \theta_t)$ is a production function of the household that depends on the level of capital stock, K_t, and on the total factor productivity (TFP), θ_t. Marginal returns are decreasing in capital $f'(K_t, \theta_t) < 0$. For simplicity, we assume that depreciation is complete, i.e. the microenterprise decides on a new optimal capital stock each period.[5] We assume that the household can also borrow at the interest rate r,and invest all borrowing capital in production. Let $r^t\left(A - A_{K_0}\right) - \sum_{i=1}^{t-1} r^{t-i} A_{K_i}$ represent net earnings on wealth after deducting borrowing discharges.

The household's decision making process consists of two steps. First, the household chooses the current level of the capital stock, and second, it chooses the current level of consumption taking into account that current consumption affects future consumption possibilities through investments. The expected life-time utility and the household's problem can be written as

$$V(K_t, \theta_t) = \max_{C_t}[u(C_t) + \beta E_t V(K_{t+1}, \theta_{t+1})] \qquad (2)$$

subject to $C_t = f(K_t, \theta_t) - K_{t+1} + r^t\left(A - A_{K_0}\right) - \sum_{i=1}^{t-1} r^{t-i} A_{K_i}$.

For brevity let $\hat{A}_t = r^t\left(A - A_{K_0}\right) - \sum_{i=1}^{t-1} r^{t-i} A_{K_i}$. The Bellman equation (2) effectively reduces the multi-period optimization to a two-stage problem. Substituting the constraint on C_t into the objective function of the Bellman equation leads to

$$V(K_t, \theta_t) = \max_{C_t}\left[u\left(f(K_t, \theta_t) - K_{t+1} + \hat{A}_t\right) + \beta E_t V(K_{t+1}, \theta_{t+1})\right]$$

The first-order condition for K_{t+1} is

[5]Note that this simplification can be rationalized by the fact that we consider an infinite number of periods.

$$u'(C_t) = \beta E_t V'(K_{t+1}, \theta_{t+1}).$$

The envelope theorem gives $V'(K_t, \theta_t) = u'(C_t) f'(K_t, \theta_t)$, which can be shifted forward one time period $E_t V'(K_{t+1}, \theta_{t+1}) = E_t u'(C_{t+1}) f'(K_{t+1}, \theta_{t+1})$, so that we get the Euler condition

$$u'(C_t) = \beta E_t u'(C_{t+1}) f'(K_{t+1}, \theta_{t+1}), \qquad (3)$$

$$f'(K_{t+1}, \theta_{t+1}) = \frac{u'(C_t)}{\beta E_t u'(C_{t+1})}.$$

This Euler condition (3) with the budget constraint (1), the initial value of capital stock K_0 and the transversality condition $\lim_{t\to\infty} \beta^t E_t u'(C_t) f'(K_t, \theta_t) K_t = 0$ form the equilibrium for the microenterprise without credit and risk constraints. The transversality condition requires that the discounted value of the limiting capital stock converges to zero, otherwise it may become infinitely large.

With credit and risk constraints

The household's problem with credit and risk constraints has the following form:

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t u(C_t) \right] \to \max_{\{C_t\}_{t=0}^{\infty}}$$

subject to $C_t = Y_t - K_{t+1} + r^t (A - A_{K_0}) - \sum_{i=1}^{t-1} r^{t-i} A_{K_i}$,

$$Y_t = f(K_t, \theta_t) + \varepsilon h(K_t, \theta_t) \quad (4)$$

$$K_{t+1} \leq B_t + A_{K_t} + n w_t \qquad (5)$$

$$B_t \leq \bar{B} \qquad (6)$$

$$A_{K_t} \leq A \qquad (7)$$

where equations (5)-(7) represent credit constraints. Production is risky according to Just and Pope (1978) where ε is a random variable with mean 0 and variance σ_ε^2. The function $f(\cdot)$ represents the expected output level, while $\sigma_\varepsilon^2 h^2(\cdot)$ is the variance of output (with the restriction $h(\cdot) > 0, \forall K_t, \theta_t$). As Just and Pope (1978) indicate, an adequate specification of risky production should include two separate functions defining the effects on the mean and on the variance of output. Our linear specification hence implies that the effect of inputs on outputs is not tied to the effects of inputs on output variability. The signs of the two effects – on the mean and on the variance – are not determined a priori even when $h(\cdot)$ follows one of the traditional production functions.

The Euler condition of the household with risk and credit constraints is

$$u'(C_t) = \beta E_t u'(C_{t+1}) f'(K_{t+1}, \theta_{t+1}) + \beta h'(K_{t+1}, \theta_{t+1}) \cdot cov(u'(C_{t+1}), \varepsilon) - \lambda \quad (8)$$

that can be rewritten as

$$f'(K_{t+1}, \theta_{t+1}) = \frac{u'(C_t)}{\beta E_t u'(C_{t+1})} + \frac{\lambda}{\beta E_t u'(C_{t+1})} - h'(K_{t+1}, \theta_{t+1}) \cdot \frac{cov(u'(C_{t+1}), \varepsilon)}{E_t u'(C_{t+1})} \quad (9)$$

in order to make explicit the marginal returns to capital.

Proposition 1. *Marginal returns to capital are higher due to credit and risk constraints.*

Proof of Proposition 1. Equation (9) allows us to disentangle the effects of credit and risk constraints to higher marginal returns to capital. The second term of the right-hand side of expression (9) indicates an increase in the costs of using capital due to credit constraints where λ is the Lagrange multiplier on condition (5), and is a measure of how tightly overall credit constraints bind. Credit constraints without risk lead to lower investments in production and are hence associated with higher marginal returns to capital.

The third term of the right-hand side in (9) represents the implications of risk. If the household is risk neutral, then $E_t u'(C_{t+1})$ would be constant so that the last term in (9) is equal to zero (because the covariance is zero), and without credit constraints the marginal productivity equals $\frac{u\prime(C_t)}{\beta E_t u\prime(C_{t+1})}$, which is the usual efficiency condition for production (3). If the household is risk averse, then the distortion due to risk aversion is defined by $h'(K_{t+1}, \theta_{t+1}) \cdot \frac{cov(u\prime(C_{t+1}),\varepsilon)}{\beta E_t u\prime(C_{t+1})}$. The sign of $h'(K_{t+1}, \theta_{t+1})$ indicates whether risk is increasing (> 0) or decreasing (< 0) in the capital stock. We assume that risk is decreasing in capital, i.e. larger firms are less responsive to risk. Furthermore, consumption increases with ε, $cov(u'(C_{t+1}), \varepsilon) > 0$. Because $E_t u'(C_{t+1}) > 0$, the second and third terms of the right-hand side of (9) are positive, that leads to higher returns to capital and lower capital accumulation due to credit and risk constraints compared with the case without any constraints. In order to derive an explicit equation for capital accumulation, we assume standard specifications of utility and production functions.

Proposition 2. *Let the household have CRRA preferences:* $u(C_t) = \frac{C^{1-\gamma}}{1-\gamma}$, *and production described by a Cobb-Douglas form* $f(K_t, \theta_t) = e^{\theta_t} K_t^\alpha$ *and* $h(K_t, \theta_t) = e^{\theta_t} / K_t^\alpha$ *(risk-decreasing):*[6] *Then, without credit and risk constraints a closed form solution will be* $K_{t+1} = \alpha\beta e^{\theta_t} K_t^\alpha$ *and* $C_t = (1 - \alpha\beta) e^{\theta_t} K_t^\alpha + \hat{A}_t$. *In the case where credit and risk constraints are present, the solution is more complex:*

$$K_{t+1} = e^{\theta_t} K_t^\alpha + \varepsilon \frac{e^{\theta_t}}{K_t^\alpha} + \hat{A}_t - \left(\frac{\alpha\beta}{\pi_k} - 1\right)^{\frac{1}{\gamma}} \left(\frac{1}{\lambda}\right)^{\frac{1}{\gamma}} \qquad (10)$$

where $\hat{A}_t = r^t (A - A_{K_0}) - \sum_{i=1}^{t-1} r^{t-i} A_{K_i}$ *and* $K_{t+1} = \pi_K e^{\theta_t} K_t^\alpha$.

Proof of Proposition 2. See Appendix.

The first term of (10) shows the positive effect of productivity on accumulation. The second term captures the effects of risk as possible deviation from expected output. The term also indicates that MEs can reduce risk by increasing K_t. The third term demonstrates that higher initial wealth allows for faster capital accumulation by relaxing credit constraints. Equation (10) also explicitly shows that a more risk averse household (high γ[7]) will increase capital investments at the expense of consumption (the last interaction term of (10) is lower if γ is

[6] In the case of risk-increasing capital, the main results hold too.

[7] In the case of CRRA utility $1/\gamma$ is the intertemporal elasticity of substitution.

higher). As entrepreneurs exhibit higher tolerance to risk (Ahn, 2010) (low γ), they tend to invest less capital.

Proposition 3. Assuming CRRA utility and a Cobb-Douglas production function, the rate of capital accumulation is higher in MEs with higher productivity, higher initial wealth and less risk exposure.

The proof of proposition 3 is fairly straightforward.

The last term in (10) measures how credit constraints interact with risk and how this interaction affects capital accumulation. From our model it can be found that

$$\frac{K_{t+1}^{withoutconstr}}{K_{t+1}^{withconstr}} = 1 + \lambda(C_t)^{\gamma}. \quad (11)$$

Equation (11) explicitly derives the gap between the constrained and unconstrained capital stock. More risk averse microentrepreneurs value foregone investment due to credit constraints less (the shadow value of risky profits is lower for them). This implies that their incentives for additional investments out of withheld profits are also lower, and, ceteris paribus, their pace of capital accumulation is lower.

Corollary 1. Assuming CRRA utility and a Cobb-Douglas production function, credit and risk constraints lead to a lower pace of capital accumulation.[8]

In addition, the microenterprise age is important for capital accumulation.

Proposition 4. The effects of credit and risk constraints on capital accumulation decrease with the age of the microenterprise.

While the microenterprise becomes older the output shocks ε sum up and the effect of risk will converge to the mean times the microenterprise's age. Hence, while the age increases, the effect of output-risk converges to 0. Indeed, new information or learning may decrease uncertainty (Jovanovic, 1982). The risk premium will then decrease over time and accumulation will (ceteris paribus) be faster in older firms.

The enterprise age also relaxes credit constraints. Older firms have higher interests on endowments. This can be seen from the term $r^T A$ in equation (10). Similarly, entrepreneurs adjust their capital stock over time and can overcome credit constraints by retaining earnings (Evans and Jovanovich, 1989; Cabral and Matá, 2003). Capital stocks would then be increasing with enterprise age.

Of course, our small model is not able to display all possible and relevant features of the capital accumulation process in MEs. One important omission is the neglect of the value of investment options (e.g. Bloom, Bond and van Reenen, 2007; Dixit and Pindyck, 2004). We have also not allowed for risk to be a function of wealth, although wealthier individuals have been shown to be less risk-averse. Furthermore, if firm owners face credit constraints and investments are partly irreversible they may have strong motives for precautionary savings. Only when the level of precautionary savings is deemed comfortable enough, risky

[8] See the proof of Corollary 1 in the appendix.

investments will be undertaken (Fafchamps and Pender, 1997; Fafchamps, 1999). Interactions between wealth and risk may also manifest themselves in (endogenous) time-inconsistent preferences. Very high discount rates are typically observed for low-wealth individuals in high-risk environments (Pender, 1996; Yesuf and Bluffstone, 2008).

Despite these omissions our model captures some of the fundamental determinants of capital accumulation and the above considerations lead to the following hypotheses to be tested subsequently: (i) returns to capital are higher in MEs, if these firms operate under credit constraints and risk; (ii) credit constraints and risk reduce initial capital stocks and result in a slower pace of capital accumulation; (iii) the impact of both risk and credit constraints on accumulation decreases with higher enterprise age.

4. Data and ME characteristics

We use data from the nationally representative Peruvian household survey (ENAHO) collected by the National Institute of Statistics and Informatics (INEI) between 2002 and 2006. The ENAHO comprises around 20,000 households each year including a (fixed, not rotating) panel sub-sample, ENAHO household panel hereafter, of about 5400 households (again nationally representative) (for details of the household panel see Table A1 in the appendix).[9] What makes the ENAHO dataset unique is that it allows constructing a panel of MEs from the household panel data.

The survey provides detailed information on individual socio-demographic and employment characteristics. Individuals who are identified as independent workers or as employers (in principal or secondary employment), and who are not working in agriculture, livestock production, or forestry, are interviewed in an Informal Sector Module.[10] This module captures the characteristics of the entrepreneurs and their production unit with up to ten employees. It also contains detailed information on input use and sales and the legal status of the firm as well as characteristics of employed workers. Firms with more than 10 employees are not interviewed. We restrict our analysis to households and MEs in urban areas, i.e. in cities with at least 4,000 inhabitants. Furthermore, in our sample of MEs we exclude those observations with either missing or non-positive values for the variables used (value-added, labor and capital), as well as MEs in the primary sector. Hence, we reduce the original dataset, and our final dataset for estimation is an unbalanced panel of MEs with 1,006, 439, 283, and 135 MEs, which we observe in two, three, four and five years, respectively. The excluded remaining 6097 firms are observed during only one year.

[9] In 2002, the survey took place during the 4th quarter (October to December). Starting from May 2003, the survey is permanent (the whole sample is distributed monthly along the year). Around 18 percent of the visited households are not interviewed as the household refuses, is absent, the house is unoccupied or other reasons (miscellaneous category). The whole unbalanced panel comprises 1435, 1153, 1870 and 2096 households being observed in two, three, four and five years, respectively. The fact that this number is increasing reflects increased effort by INEI to create a larger panel dataset. Quite a number of panel households were not interviewed in consecutive years.

[10] In urban areas, 10 percent of self-employed individuals work in agriculture, livestock production, or forestry, and are consequently not included in the Informal Sector Module.

Hereafter we refer to the sample that we use for most analyses below as the 'ENAHO firm panel'.

Table 1 shows the employment structure of the labor force in the ENAHO household panel. The importance of the informal sector is remarkable, even compared to Latin American averages. More than a third of the Peruvian workforce is self-employed. Another almost 40 percent work as paid worker in informal firms, defined as firms without registration or written accounts, or as unpaid family workers. Compared to men, women are less likely to be wage-employed and more often engaged in unpaid family work.

Table 1: Structure of urban employment in Peru (in percent)

	Total	Male	Female
Wage employment	48.70	53.25	43.05
… of which in informal firms (%)	*39.96*	*41.56*	*36.27*
Self-employment or employer	36.16	34.28	38.49
Unpaid family work	7.99	5.54	11.04
Unemployed	7.15	6.99	7.36
Total	100	100	100

Source: Authors' calculation based on the ENAHO household panel.
Notes: The sample contains in total 27,249 urban workers, of which 15,096 are men and 12,153 are women.

Some basic characteristics of Peruvian MEs are highlighted in Table 2 using the ENAHO firm panel.[11] MEs are typically very small with a mean firm size of 1.6 including the owner. Almost one third of the firms operate with the help of unpaid family members and only 12 percent employ paid staff. The average enterprise age of about 8 years is higher than one might probably expect. The ME owner is, on average, 42 years old and has been to school for 10 years. Almost half of the owners are female. Incomes from MEs are fairly low, with mean and median monthly value-added of about USD 150 and 90.[12] The median capital stock of USD 80 suggests that most activities do not require much investment. On the other hand, a mean capital stock of USD 700 implies that some firms have relatively high capital stocks.[13] Most MEs can be found in *petty trading* (38 percent) followed by *transport* (15 percent), *hotels and restaurants* (13 percent), and *other manufacturing & food* (13 percent). The different industries are very heterogeneous in many characteristics, for example in capital stocks, gender-composition, or number of employees.

[11]Descriptive statistics hardly differ between the ENAHO firm panel and a dataset that uses all MEs from the pooled cross-sectional data from 2002 to 2006 (as opposed to using only those of the household panel). Detailed summary statistics are available from the authors upon request.

[12] All monetary values are converted into 2001 USD. See Notes of Table 2 for details. Value-added is the value of monthly sales plus self-consumed production minus expenses other than on labor and capital (i.e. expenses for intermediate inputs and electricity).

[13]Here, and in all the below analyses capital stocks are computed as the total replacement value of machines, furniture, vehicles, and tools used by the production unit. We exclude the value of land and buildings.

Table 2: Descriptive statistics of MEs by industry

Activity	All	M.& food	Constr.	W./R. shops	Petty trad.	H.& R.	Tran.	Oth. serv.
ME characteristics								
Employees	0.63	0.82	0.65	1.12	0.68	0.91	0.13	0.45
Self-employed (%)	63.2	59.8	64.5	45.2	55.9	46.2	91.0	78.3
With paid employees (%)	12.1	21.3	31.5	36.7	5.8	12.2	7.7	11.1
With unpaid employees (%)	31.4	32.7	20.5	39.5	41.2	49.7	4.1	16.4
Monthly value-added (mean)	153	148	166	305	134	135	165	168
Monthly value-added (median)	90	81	129	190	72	74	133	76
Capital (mean)	697	823	127	962	217	192	2396	691
Capital (median)	83	115	23	153	41	65	1610	111
Capital (median) (new MEs)	58	65	26	227	29	46	1251	37
Enterprise age	7.9	11.7	11.6	9.7	7.8	6.0	5.7	6.4
Entry rate (%)	29.4	30.4	37.9	24.0	25.4	33.8	26.9	39.0
Exit rate (%)	32.0	34.1	41.6	25.8	26.6	33.0	31.9	46.0
Owner's characteristics								
Gender: Male (%)	52.1	55.5	99.8	84.0	31.7	12.3	99.8	64.6
Owner's age	41.9	44.7	42.5	40.0	43.3	43.9	37.2	38.2
Owner's years of schooling	10.5	10.5	10.4	11.3	9.8	9.1	11.2	12.7
Observations	7960	1026	454	281	2995	1051	1194	959

Source: Authors' calculation based on the ENAHO firm panel.
Notes: Sectors are other manufacturing & food, construction, wholesale/retail shops, petty trading, hotels and restaurants, transport, and other services. Monetary values are in constant Dec. 2001 Nuevo Sols (using the INEI Consumer Price Index) and converted into US-$ using the Dec. 2001 nominal exchange rate.

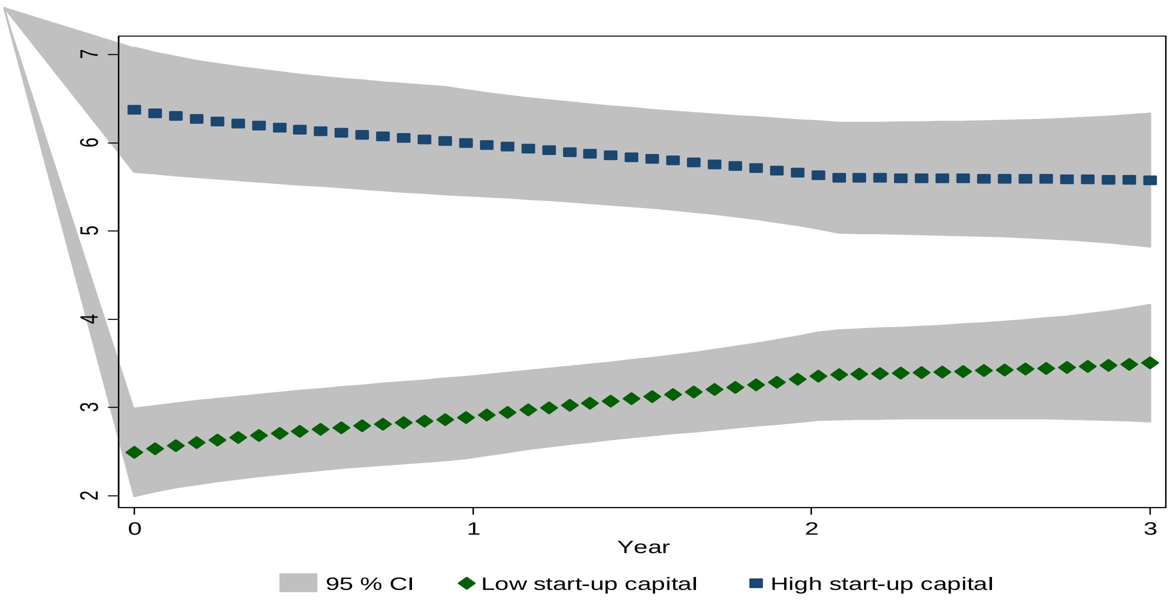
Data presented in Table 2 show that MEs are being established and closed at a substantial rate. Depending on the industry, new starts during the year account for 24-38 percent of all existing firms. Similarly, of all observed firms between 26 and 42 percent close every year.[14] We identify the firms by their owner. If a business is handed over from the owner to his wife, for example, this will be counted as business closure and opening. Consequently, our entry and exit rates may only be taken as an upper bound. However, these numbers are similar to those in Mead and Liedholm (1998) who find entry and exit rates in Sub-Sahara Africa ranging from 20-32 percent. Entry (as well as exit) rates in our sample are higher in sectors that require less capital. New firms start with a 30 percent lower median capital stock on average. Capital accumulation is quite heterogeneous as well; for example, MEs in *wholesale/retail shops* – on average – even reduce their capital stock (see Table 1).

Figure 1 provides another indication that Peruvian MEs tend to adjust their capital stocks over time. Based on a sample of firms with a maximum age of three years when first observed, this figure shows the capital stock (year effects controlled) of MEs that can be found in the first and fourth quartile of the distribution of initial capital stocks,

[14] Entry and exit rates are computed using only the panel households. Entry (exit) rates in this study are computed by dividing all new (exiting) firms appearing in a given year by the number of firms existing at the beginning of that year. Exiting firms are no longer observed in the subsequent year.

respectively.[15] The initial difference in capital stocks between these two groups is very large. Firms in the fourth quartile start with a log capital stock that is almost four times higher than from those in the first quartile. This implies that the capital stock is more than 50 times higher in this group. Yet, MEs that start with lower capital stocks accumulate, while those with a high capital stock even tend to reduce their capital stocks. After 3 years, capital stocks in the two groups still differ significantly. Nevertheless, the log difference has reduced to about 2.5 implying that capital stocks in firms with high initial capital stocks are only about ten times larger than in the group that starts with low levels.

Figure 1: Capital accumulation and de-cumulation of young firms over time (sub-sample)



Source: Authors' calculation based on the ENAHO firm panel.
Notes: The figure is based on a sub-sample including only young firms with a maximum age of three years old when first observed and that are observed in at least four periods. These are 168 firms.

5. Empirical analysis

5.1 Capital returns

We first estimate a production function to test whether we can find high returns to capital in Peruvian MEs. For estimation purposes we take a log-linear transformation of the production function:

$$y_{it} = \alpha_0 + \beta_l l_{it} + \beta_k k_{it} + u_{it} \quad (12)$$

$$u_{it} = \omega_{it}^* + \varepsilon_{it} = \lambda_i + \omega_{it} + \gamma_t + \varepsilon_{it},$$

[15] More precisely, we follow Banerjee and Munshi (2004) by regressing the logarithm of capital on a full set of year dummies. We then plot the residual of this regression plus the mean value against time using local polynomial smoothing based on an epanechnikov kernel.

with lower case letters referring to log values of value-added, labor, and capital of firm *i* at time *t*, respectively. The error term u_{it} may be decomposed into two components: ε_{it}, an i.i.d. measurement error term and $\omega^*_{it} = \ln(\Omega_{it})$, a productivity term. Total factor productivity (TFP) is split into an element common to all firms (γ_t) and firm-specific element, which is divided into a time-invariant component (λ_i) and a time-variant productivity shock (ω_{it}). We estimate the production function for the whole sample of firms and for each industry. We, first, employ a random effects (RE) panel model, which explicitly includes time-variant and time-invariant components. A RE regression gives unbiased estimates only if the individual effect ω_{it} influences value-added directly. However, it may also simultaneously determine input factor use which may bias our coefficients. Indeed, the main problem for estimation of (12) using panel data arises from this `transmission bias' when an econometrician cannot efficiently disentangle the effect of inputs from productivity shocks[16]. Several solutions for the endogeneity of input choices with regard to unobserved productivity have been proposed for large *N*, small *T* samples, including structural and GMM estimators. The latter rules out dynamics in inputs and productivity and proves to be most pertinent for tackling the transmission bias problem. However, it reduces our sample by about 50 percent as it requires at least three time periods. In order to address the problem of high persistence in the data and serial correlation in unobservables, the approach assumes a first order autoregressive process AR(1) for ω_{it}:

$$\omega_{it} = \rho\omega_{it-1} + \xi_{it}, \qquad |\rho| < 1, \qquad \xi_{it} \sim MA(0)$$

Deriving ω_{it} from (12) and replacing it in the AR(1) process, we get the dynamic regression model with the lagged levels of the dependent and independent variables on the right-hand side:

$$\begin{aligned} y_{it} &= \rho y_{it-1} + \beta_l l_{it} - \rho\beta_l l_{it-1} + \beta_k k_{it} - \rho\beta_k k_{it-1} + \\ &(1-\rho)(\alpha_0 + \lambda_i) + (\gamma_t - \rho\gamma_{t-1}) + \xi_{it} \end{aligned} \tag{13}$$

and plus $(\varepsilon_{it} - \rho\varepsilon_{it-1})$ that disappears if there is no measurement error. We estimate (13) by Blundell and Bond (1998) system GMM (BB GMM). To cope with the `transmission bias' we use past values of the regressors as instruments. In all regressions we also add year dummies to account for the effect of common shocks. The unobserved productivity term $\omega_{it}$ is then no longer a problem. Equation (13) implies `common factor restrictions', as the coefficient before l_{it-1} is required to be equal to the negative product of the coefficients before y_{it-1} and l_{it}, and the coefficient before k_{it-1} is required to be equal to the negative product of the coefficients before y_{it-1} and k_{it}. We test implied restrictions, and if not rejected, we can employ a minimum distance estimator and the Delta method to obtain standard errors.[17] If common factor restrictions are rejected, then the long-run coefficients

[16] For more details on the `transmission bias' and other difficulties for production function estimators see Eberhardt and Helmers (2010).

[17] This is easily implemented in Stata using the COMFAC command thanks to Mans Söderbom's code.

for labor and capital can be computed as nonlinear combinations of the estimated coefficients $\frac{\widehat{\beta_l}-\widehat{\rho\beta_l}}{1-\widehat{\rho}}$ and $\frac{\widehat{\beta_k}-\widehat{\rho\beta_k}}{1-\widehat{\rho}}$ (Eberhardt and Helmers, 2010).

Moreover, the semi-parametric proxy estimator proposed by Levinsohn and Petrin (2003) (henceforth LP) is used as a further robustness check. This approach assumes an introduction of term $g(\cdot)$ into (12), which is typically a lagged polynomial of a flexible input and capital. This flexible input is supposed to respond freely to the current productivity shock and capital level so it can be used as a proxy for the productivity shock. If $g(\cdot)$ turns out an adequate proxy, then the compound error term in (12) comprises productivity innovation, common time-specific shocks, and measurement error. Our LP estimator uses within-plant changes in intermediate/electricity inputs as a proxy for productivity shocks. Unfortunately, about 15 percent (66 percent) of all MEs in our sample do not use intermediate (electricity) inputs, and would be excluded from the analysis, hence we do not give too much importance to this estimator. It is only used as a robustness check because of the sample reduction problem. Finally, the FE estimator uses only the variation within firms over time. It is unbiased under the assumption of the unobserved firm-specific productivity being time-invariant. Using only the within variation, coefficients reflect marginal productivities which implies that potential start-up costs are ignored. Yet, the FE estimator suffers from `Nickell bias' (Nickell, 1981), which leads to downward biased coefficients if T is small.

Tables 3 and 4 present the production function estimates as well as the implied marginal returns to capital based on the RE and GMM estimations. The RE coefficients are very much in line with previous studies with a log labor coefficient in the range between 0.5 and 0.8, and a log capital coefficient between 0.1 and 0.2.[18] There is quite some sectoral variation, in particular in the capital coefficient. The differences in the sector specific R2s also suggest that there is a lot of heterogeneity in the within sector variance of returns, probably partly driven by differences in the extent of reporting errors. The output-capital elasticity is highest in manufacturing (0.22) and lowest in petty trading (0.08). In the most capital intensive sector *transport* the returns are rather low with only 2 percent, while they amount to 80 percent in the least capital intensive sector *construction*. If an entrepreneur invests an additional USD 10 USD into her ME, monthly income increases by USD 6.8 in *construction*, and only by USD 0.2 in *transport*. The average monthly returns to capital are with 14 percent rather high. To relate these results on marginal returns to income levels of the entrepreneurs: At mean value-added and capital stock these marginal returns would

[18] Westbrook and Tybout (1993), for example, analyze returns to scale in firms that are observed in up to 8 years and have at least 10 employees in Chile. They find OLS log labor and log capital coefficients in most industries to be in the range between 0.7 and 0.9, and 0.2 and 0.3, respectively. Their within-plant estimates are in the range between 0.4 and 0.6, and 0.1 and 0.3, respectively. The authors argue that their OLS estimates appear to suffer from an upward bias while their within-plant estimates are likely to be downward biased, and identify measurement error in capital stocks to drive the latter result. This problem is certainly acerbated in our dataset covering microenterprises with fewer periods and up to 10 employees.

correspond to a permanent income gain of 6 percent in *construction*, but only 0.1 percent in *transport*.

Table 3: Production function estimates for Peruvian industries (RE)

Sectors	All	M.& food	Constr.	W./R. shops	Petty trad.	H.& R.	Tran.	Oth. serv.
Log labor	0.605***	0.619***	0.642***	0.697***	0.582***	0.690***	0.486***	0.558***
SE	(0.015)	(0.036)	(0.045)	(0.066)	(0.028)	(0.041)	(0.040)	(0.040)
Log capital	0.118***	0.220***	0.141***	0.104***	0.080***	0.076***	0.139***	0.142***
SE	(0.005)	(0.015)	(0.026)	(0.029)	(0.011)	(0.020)	(0.014)	(0.016)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Outlier$^+$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Average K	84	87	23	150	39	57	1115	83
Average Y	101	78	127	228	93	101	150	76
Average Y/K	8.8	7.2	16.1	6.2	12.2	5.3	1.7	10.1
M. returns	0.143	0.198	0.774	0.159	0.191	0.136	0.019	0.131
R2 overall	0.39	0.57	0.47	0.49	0.25	0.45	0.34	0.38
N	7857	1012	444	274	2952	1039	1170	944

Source: Authors' calculation based on the ENAHO firm panel.

Notes: Robust standard errors (SE) in parenthesis * p<0.10, ** p<0.05, *** p<0.01. $^+$ In all regressions, we drop influential outliers from our sample (and sub-samples) that we identify by the DFITS-statistic obtained from a precedent OLS regression (See Belsley, Kuh, and Welsch, 1980). We use a cut-off-value $|DFITS| \quad = 3\sqrt{\mathrm{k/N}}$, with k, the degrees of freedom (plus 1) and N, the number of observations. This reduces the sample by 1-2 percent. Marginal returns to capital $\frac{dY}{dK} = \gamma * \frac{Y}{K}$ are evaluated at Y=exp(mean(log Y)) and K=exp(mean(log K)).

Regarding the BB system GMM estimation we implement several specifications. We begin with lagged labor and capital as endogenous and profit, capital stock and labor as predetermined following the assumptions for the structural (LP) estimator according to Eberhardt and Helmers (2010). Then, we vary these assumptions and the lag length and check the validity of instruments using the overidentification tests (Hansen, Sargan, Difference Sargan). We use the asymptotically more efficient two-step procedure and correct for possible small sample downward bias as suggested by Windmeijer (2005). In all GMM models we specify year dummies to account for common year shocks. Table 4 presents the GMM results based on the most preferred specification based on diagnostic tests: the profit is predetermined, and capital and labor each lagged by one period are used as instruments[19]. For the whole sample the dynamic specification is well supported: the hypothesis of no first order serial correlation (AR1) is boundary rejected and the hypothesis of no second order serial correlation (AR2) is not rejected. However, for most industries the hypotheses of no both first and second order serial correlation are not rejected.

[19]The results for the baseline GMM specification are reported in the appendix. The coefficients on labor and capital are much closer to the RE estimates but Sargan and Hansen tests are strongly rejected for most cases. The reported tests for first and second order serial correlation (AR1 and AR2 respectively) confirm the dynamic specification. The hypothesis of no first order serial correlation is rejected but the hypothesis of second order serial correlation is not rejected.

Table 4: Production function estimates for Peruvian industries (GMM)

Sector	All	M.& food	Constr.	W./R. shops	Petty trad.	H.& R.	Tran.	Oth. serv.
Log labor	0.831***	0.857***	1.006***	0.594***	0.637***	1.170***	0.851***	1.129***
SE	(0.072)	(0.129)	(0.188)	(0.155)	(0.115)	(0.185)	(0.141)	(0.249)
Log capital	0.169***	0.224***	0.188*	0.008	0.097*	0.102	0.095***	0.192**
SE	(0.026)	(0.065)	(0.121)	(0.081)	(0.053)	(0.076)	(0.036)	(0.101)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Outlier[+]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Average K	106	175	29	124	50	76	1181	99
Average Y	122	112	161	239	99	139	159	118
Average Y/K	8.8	4.2	24.9	11.5	12.6	6.7	0.5	4.4
M. returns	0.194	0.143	1.026	0.016	0.193	0.186	0.013	0.228
N	2244	284	114	78	943	310	345	169
N of groups	856	132	50	40	404	141	142	83
Sargan P-value	0.001	0.220	0.153	0.001	0.231	0.244	0.000	0.417
Hansen P-value	0.030	0.416	0.374	0.299	0.752	0.757	0.215	0.768
AR(1)	0.101	0.433	0.876	0.842	0.231	0.313	0.236	0.096
AR(2)	0.797	0.957	0.491	0.516	0.981	0.942	0.840	0.256
COMFAC	0.372	0.649	0.998	0.817	0.408	0.908	0.066	0.659

Source: Authors' calculation based on the ENAHO firm panel.

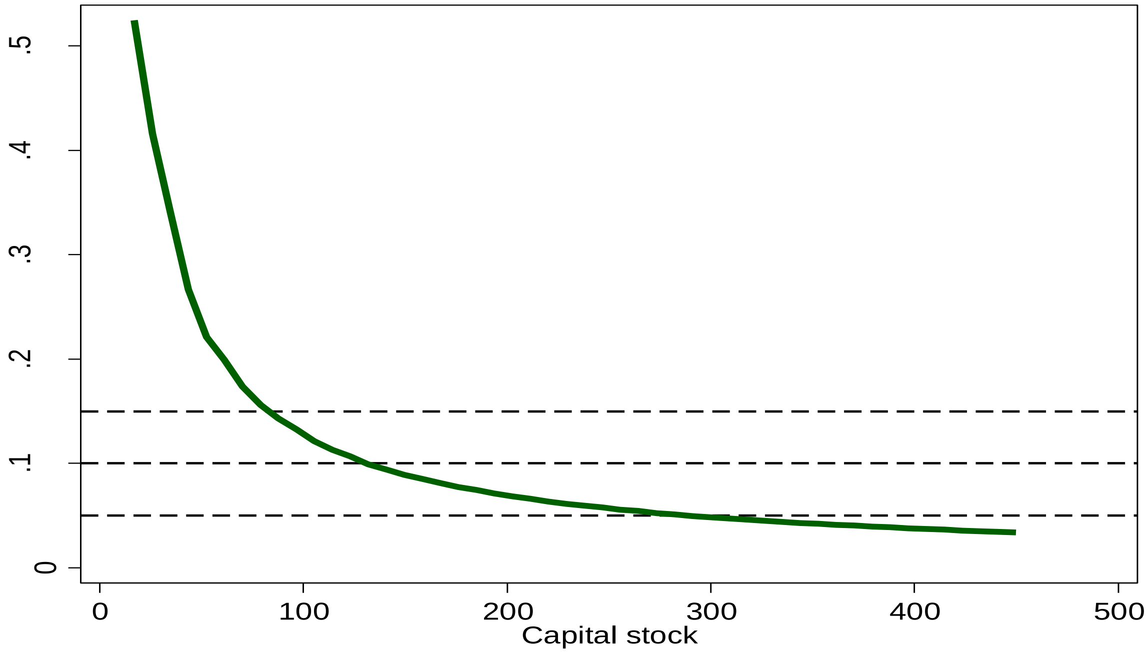
Notes: Influential outliers are dropped as above. Marginal returns to capital $\frac{dY}{dK} = \gamma * \frac{Y}{K}$ are evaluated at Y=exp(mean(log Y)) and K=exp(mean(log K)). The estimator employed is BB – Blundell and Bond (1998) System GMM with predetermined profit and lagged labor and capital as instruments. We use the 2-step estimator with the Windmeijer (2005) correction. The sample includes all entrepreneurs observed during at least three time periods. COMFAC is a minimum distance test of the common factor restrictions. P-values reported for all test statistics. For the BB estimator the Difference-in-Sargan and Hansen statistics test the validity of all the additional instruments employed for the levels equation. All long-run coefficients are computed using the Minimum Distance estimator (code from Mans Söderbom) and employing the Delta method to obtain standard errors (absolute values reported).Robust standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01.

The coefficients on labor are much higher than the RE estimates, however remaining close to the expected magnitude of 0.83 for the whole sample and for some other industries between 0.59 and 0.86. Construction, hotels/restaurants and other services report the highest labor coefficient with an order of magnitude of more than 1. This is not unusual in such estimates and may be due to weak instruments or relatively high correlation between inputs. The coefficient associated with the lagged dependent variable is as expected around 0.5. The coefficients on capital are in line with theory and previous RE estimates between 0.1 and 0.22. Only for wholesale and retail shops the coefficient on capital is very low and insignificant that may be caused by a very small sample size. The Sargan and Hansen tests of overall instrument/moment restriction validity for these models indicate that the null hypothesis of exogeneity is not rejected in most cases. The models do not suffer from serial correlation given that AR(1) is introduced by construction. The 'common factor restrictions' are not rejected except for transport where they are not weakly rejected at 5%. The long-run coefficient on capital for *transport* is 0.17, which is close to that for the whole sample.

The GMM estimates also show a high heterogeneity in within and between sector variance of returns to capital. The marginal returns to capital are considerably high, almost 20 percent on average. The lowest returns of only 2 percent are observed in *transport* and *wholesale and retail shops*. In the least capital intensive sector, c*onstruction,* the marginal returns of capital are highest and come to 100 percent. The pattern of these estimates is very similar to the RE estimates reported above.

Figure 2 presents the returns to capital as a function of the capital stock based on the estimates presented in Table 3 (for all industries). Except for the capital stock, we take mean values in all other variables. As previous studies, our analysis of returns to capital in Peruvian MEs shows very high marginal returns to capital that decline rapidly as firms accumulate capital. Yet, in a range up to 130 USD of capital stock marginal returns are well above 10 percent monthly. More than 57 percent of MEs can be found in this range and may hence be able to realize these returns if provided the necessary capital. As laid out above, the lack of credit may, however, not be the single reason for low capital stocks and unexploited opportunities. Risk may also cause lower investment and hence lower capital stocks.

Figure 2: Marginal returns for firms with an interquartile capital stock (p75 - p25)



Source: Authors' calculation based on the ENAHO firm panel.
Notes: Marginal returns at lower than p25 capital stocks are very high which would imply a wide y-axis scaling. In contrast, they are close to zero for higher capital stocks, and a maximum value of 32,000 USD would imply a wide x-axis scaling. For a better graphical understanding, we choose to display only the interquartile range. We predict the profit level given the capital stock and mean values in all other variables. Whereas the mean log labor is 5.34 for the whole sample, it is only 4.94 for firms with a capital stock in the first quartile of capital stock. This results in a slight overestimation the returns to capital at low levels. Returns to capital are proxied by (ΔY/ΔK) with ΔX = Xt – Xt-5. Results are robust to an extension of the lag length to 10 or 20.

5.2 Capital accumulation

In this section, we attempt to identify the factors that restrict capital accumulation. In line with our theoretical framework, we first examine the determinants of the start-up capital stock. The investment equation then takes the following functional form:

$$k_{i0} = \mu_1 + \mu_2\Omega_{it}^{e} + \mu_3 W_{i0} + \mu_4 R_{i0} + \mu_5 R_{i0} W_{i0} + e_{i0}, \tag{14}$$

In a second step, we analyze the accumulation process:

$$\Delta k_{it} = \rho_0 k_{it} + \rho_1 + \rho_2 \Omega_{it} + \rho_3 W_{if} + \rho_4 R_{it} + \rho_5 R_{it} W_{if} + u_{it}. \tag{15}$$

All variables are defined as before, e_{i0} is an error term, and $u_{it}$is the combined error term. Capital growth Δk_{it} is measured as the change in log capital ($k_{it+1} - k_{it}$). The term $R_{it}W_{if}$ is the interaction between risk and wealth (from the first observed period in business). We now briefly discuss the key right-hand-side variables and how we proxy them: credit constraints, productivity, and risk. Their distributions as well as means and standard deviations are presented in Figure A1 in the appendix.

The *expected productivity* is proxied by owner's age and its square, years of education, and a dummy for male entrepreneurs. In equation (15), *productivity* Ω_{it} is measured as the time invariant residual λ_i obtained by either a RE or GMM estimation of the production function. As we take the predicted values of Ω_{it} from the production function estimations in Tables 3 and 4, we use bootstrapping to obtain a consistent estimate of the standard errors in (14) and (15).

Credit constraints are proxied by the wealth level W_{if} of the household, as household assets may serve as collateral for credit. The household wealth index is derived from the first principal component of a set of indicators of ownership of household assets. The latter include only non-business assets, such as color televisions and the condition of the house, for example the state of the walls and the quality of sanitary facilities. Household wealth, however, may be correlated with unobservable characteristics, such as entrepreneurial ability or motivation, which also affect the start-up capital stock and growth. This correlation is likely to be positive implying an upward bias of the estimated coefficient. This potential bias should be somewhat reduced by including the above proxies (age, education, gender) for *expected productivity* in equation (14) and the productivity measure in equation (15). Possible problems of reverse causality are addressed in equation (15) by using for the wealth index the information on asset ownership provided in the first observed period in business (W_{if}).

Risk is difficult to operationalize. Recent empirical contributions interested in the effects of risk on ME performance have therefore typically relied on measuring risk aversion that will act as a lever for the effects of risk.[20] Such information is not available in the standard type of survey used in our analysis. Yet, the much bigger sample size and the panel dimension of the Peruvian data – as compared to more specialized datasets on MEs – allow constructing a series of proxies that rely on the variation of sales both between and within firms as well as

[20] De Mel et al. (2008) illicit risk aversion using experimental approaches,

involuntary exit of entrepreneurs. We compute measures of risk at the sector and firm-level. At the sector level, these measures reflect market, technological and strategic risks. For example, economic sectors are likely to differ systematically as regards the volatility of product prices, the risk of investing in new technology, or the exposure to contractual risks. When computed at the firm-level, the variance of sales should additionally reflect personal risks, for example sickness of contributing household members. Yet, the previously mentioned risks (market, technological, and strategic) may also have an idiosyncratic component and firms have different capacities to mitigate them. The proposed measures are hence not able to distinguish the different sources of risk; neither are the sector- and firm-level proxies clearly conceivable as proxies of covariate/common or idiosyncratic risks, respectively.[21]

Specifically, we, first, measure risk by the variation of sales, a “classical” proxy for risk (Ritter, 1984). We compute this variation at the sector level with sectors being disaggregated as finely as possible while keeping the number of observations in each sector cell at least at 30. To this end, we use the pooled cross-sectional sample with almost four times as many observations that allows for a finer disaggregation. Such a procedure yields 85 sector cells, for which we compute the coefficients of variation in sales.

Second, we propose to measure idiosyncratic risk using the time-variant component from the RE or GMM estimation of equation (12). More specifically, we construct a measure of exposure to risk RES_i from the residual ε_{it}:

$$RES_i = \sum_{t=1}^{T} |\varepsilon_{it}| / T_i, \qquad (16)$$

where T_i refers to the number of periods, in which we observe the ME *i*.[22] This measure, based on average absolute time-variant RE or GMM residuals, reflects only firm-specific volatility in value-added observed over the lifespan of the ME. It captures the ex-ante effects of anticipated shocks and cyclical effects as well as possibly non-anticipated ex-post effects of shocks. The possible transmission bias of the RE model estimates may lead to an overestimation of the productivity level ω_{it} and, more importantly, a strong correlation with the input factors. Yet, it is difficult to say whether and how this bias affects the time-variant residuals. By definition these residuals sum up to zero for each firm. Their estimated variance may, however, be overestimated, for example if too much of possible anticipated cyclical effects now enter ε_{it}, in particular since labor coefficients are underestimated. The AR(1) approach to ω_{it} is supposed to rule out this problem. However, to guard against the importance of such effects we test whether ε_{it} is correlated with input use; and it turns out that this correlation is very low (<0.05).

Third, we can average RES_i at the sector-level (49 sector cells, if only the panel data is used). Fourth, we use the share of entrepreneurs that are found to be unemployed, helping as unpaid family worker, or even leaving the labor force in the subsequent period to construct

[21] For advanced economy settings, data may allow for distinguishing these different “levels” of risk. See for example Michelacci and Schivardi (2013) and the literature cited therein.

[22] The average T_i is 3.3.

a sector-specific involuntary exit rate (45 sector cells of the firm panel).[23] Except RES_i all risk proxies are measured at the sector level, and are constant over time. We opt for using sector-cells rather than sector-year cells to measure sector-specific risk. This permanent risk measure is constant over time, and it is likely to be more relevant than temporary risk shocks for investment decisions. Moreover, it allows a finer sector disaggregation.

We report the pairwise correlations of these risk proxies in Table A2 in the appendix. In line with expectations, all correlations are positive. It is difficult to judge which of these imperfect proxies is most adequate; and they may be capturing different dimensions of business risk. Due to the correlation between the single proxies they cannot be included jointly in a regression. From a combination of these proxies, we therefore propose to construct a risk proxy using principal components analysis (see Table A2 in the appendix). The first principal component that we will use as risk proxy in the subsequent analysis explains about 38 percent of the “total variance” in the four variables (53 percent without the RE idiosyncratic RES_i).

With our right-hand side variables specified, we now analyze the determinants of the start-up capital stock. The first columns of Tables 5 and 6 present the OLS estimation results of equation (14) based on the RE (Table 3) and GMM (Table 4) estimations of the production function, respectively. The sample is restricted to start-up MEs defined as being owned by entrepreneurs who did not have a firm in the previous year. By comparing Tables 5 and 6 we should not forget that the GMM estimation considerably reduces sample size. In line with the theoretical considerations, asset-rich entrepreneurs start with a considerably higher capital stock. An increase in the wealth level by 1 standard deviation raises the capital stock by 45 percent from 60 to 87 USD for a hypothetical ME with mean values. This effect is lower for the case of GMM production function estimates by about 15% so the increase will be from 60 to only 80 USD. Risk is associated with a lower initial capital stock.[24] Its effect is close for both RE and GMM production function estimates. If the risk exposure of our hypothetical firm decreases by 1 standard deviation, its capital stock is about 90 percent higher and amounts to about 115 USD. Furthermore, the interaction term of wealth and the risk proxy has a positive coefficient. This effect is much higher for the GMM production function estimates by about 80 percent. This positive effect of the interaction term suggests that at higher levels of wealth the negative effects of risk are mitigated. An alternative reading would be that credit constraints are less binding in risky environments where risk-averse entrepreneurs attach less value to foregone profits.

The four remaining columns of Table 5 present the results from the RE and GMM accumulation regressions (in Table 6 GMM estimation only). The most important factor appears to be the pre-investment capital stock, i.e. small MEs grow considerably faster. An

[23] Only entrepreneurs that are observed in the subsequent period can be classified as exiting or surviving, i.e. we do not have this information for the last round of the surveys and entrepreneurs in households that were dropped from the panel.

[24] Here, the risk index does not contain the idiosyncratic measure as it is only available for new firms that are observed in at least two periods. This would reflect a strong selection and reduce our sample by two thirds.

increase in the capital stock by 1 percent leads to an approximate decrease of the growth rate of 0.9 percentage points (of only 0.3 in case of RE accumulation regression). As expected, productivity also fosters accumulation: an increase of productivity by 1 standard deviation results in an increase of the growth rate by about 8 percentage points.[25] This large effect is observed in both RE and GMM estimations. Retained earnings thus appear to play a key role in fostering accumulation in productive firms. In addition, young firms grow faster with a turning point at a relatively high enterprise age of 43 years. This is in line with previous evidence and may be partly due to depreciation of a capital stock that is not being renewed or replaced (Mead and Liedholm, 1998; Faijnzilber et al., 2006). These (pure) age effects are rather small; the growth rate of an ME declines by about 4 percentage points with one additional year of existence.

The effects of credit constraints and risk on investment tend to be very large. Both coefficients are highly significant and the implied effects sizable. An increase in the wealth level by 1 standard deviation raises the growth rate by 10 (RE) or 35 (GMM) percentage points. If risk decreases by 1 standard deviation, the growth rate is about 20 percentage points higher according to the RE estimation or 40 percentage points higher ceteris paribus according to the GMM estimation.[26] The interaction term between risk and wealth has the expected positive sign, a finding consistent with more wealthy entrepreneurs being less risk averse.

Finally, we include interactions between enterprise age and the constraint proxies (last columns of Tables 5 and 6). The coefficient of the risk index alone increases and we also find evidence of this effect being reduced slowly over time for both RE and GMM estimations (according to RE it reduces to zero only after 62 years). The wealth-age interaction effects are small and only significant in the RE estimation. This may be because we already control for the effect of retained earnings, and so we do not necessarily expect a large age-wealth interaction effect.[27]

Summing up, both risk and credit constraints are key factors for explaining low initial capital stocks and slow capital accumulation in MEs. This suggests that these factors are also behind the high returns to capital that we observe at low levels of capital. The “mitigating” effect of wealth in the presence of risk may point at an important role of behavioral factors (high discount rates, lack of self-control) for investment decisions.

[25] We computed beta-coefficients for ease of interpretation. Details are available from the authors upon request.

[26] This numbers hold for firms with mean values. As wealth and risk are normalized, the interaction term rules out.

[27] In all GMM specifications Sargan and Hansen tests of overidentifying restrictions do not reject the hypotheses of the validity of estimates. Further, the tests for serial correlations support the dynamic model as the hypothesis of no first order serial correlation is rejected.

Table 5: Estimated functions of start-up capital and accumulation (based on RE production function estimates from Table 3)

Dependent variable:	*start-up k (OLS)*	*Δ k (RE)*		*Δ k (GMM)*	
Log capital		-0.457***	-0.469***	-0.895***	-0.892***
		(0.026)	(0.028)	(0.080)	(0.080)
Wealth index	0.178***	0.135***	0.160***	0.229***	0.235***
	(0.032)	(0.016)	(0.020)	(0.028)	(0.032)
Risk index	-0.492***	-0.243***	-0.313***	-0.427***	-0.501***
	(0.058)	(0.028)	(0.035)	(0.052)	(0.065)
Wealth * Risk	0.044**	0.033***	0.034***	0.031**	0.032**
	(0.024)	(0.011)	(0.012)	(0.015)	(0.015)
Enterprise age		-0.023***	-0.022**	-0.0397***	-0.036***
		(0.007)	(0.009)	(0.011)	(0.011)
Enterprise age^2		0.0005**	0.0004*	0.0003**	0.00069**
		(0.00022)	(0.00025)	(0.0004)	(0.0003)
TFP		0.208***	0.215***	0.192**	0.164**
		(0.059)	(0.053)	(0.079)	(0.077)
Enterprise age * risk			0.008***		0.0097**
			(0.003)		(0.004)
Enterprise age * wealth			-0.002*		-0.0007
			(0.002)		(0.002)
Year dummies	Yes	Yes	Yes	Yes	Yes
Outlier controlled+	Yes	Yes	Yes	Yes	Yes
N	1146	2443	2435	1820	1814
R2	0.239	0.189	0.199		
Sargan P-value				0.449	0.409
Hansen P-value				0.765	0.729

Source: Authors' calculation based on the ENAHO firm panel.

Notes: Regression of equation (14) includes: a constant, age and its square, years of education, and a dummy indicating male entrepreneurs. Robust standard errors are bootstrapped and presented in parentheses * p<0.10, ** p<0.05, *** p<0.01. In all regressions, we drop influential outliers according to the procedure explained under Tables 3 and 4. In the estimation presented in the first column, the risk index does not contain the idiosyncratic measure as it is only available for new firms that are observed in at least two periods. This would result in a strong selection and reduce our sample by two thirds.

Table 6: Estimated functions of start-up capital and accumulation (based on GMM production function estimates from Table 4)

Dependent variable:	*start-up k (OLS)*	*Δ k (GMM)*	
Log capital		-0.861***	-0.867***
		(0.084)	(0.085)
Wealth index	0.132*	0.206***	0.220***
	(0.068)	(0.029)	(0.036)
Risk index	-0.527***	-0.503***	-0.579***
	(0.134)	(0.060)	(0.074)
Wealth * Risk	0.189***	0.0391*	0.0432**
	(0.071)	(0.017)	(0.017)
Enterprise age		-0.0429***	-0.0453***
		(0.0145)	(0.014)
Enterprise age^2		0.0007	0.0008*
		(0.0004)	(0.0004)
TFP		0.181***	0.186***
		(0.047)	(0.046)
Enterprise age * risk			0.007*
			(0.004)
Enterprise age * wealth			-0.00085
			(0.0024)
Year dummies	Yes	Yes	Yes
Outlier controlled+	Yes	Yes	Yes
N	136	1218	1210
N of groups		748	744
R2	0.280		
Sargan P-value		0.156	0.247
Hansen P-value		0.522	0.619

Source: Authors' calculation based on the ENAHO firm panel.
Notes: VA and capital are predetermined.

5.3 Robustness

We first provide some robustness checks for our production function estimates and the corresponding estimated returns to capital. Then, we examine the robustness of the results on accumulation with a focus on the effects of risk.

Table 7 presents the LP and FE estimates of equation (12). The LP coefficients on capital are very similar to the RE coefficients, but the labor coefficients are significantly lower. This gives us some confidence that the RE estimates on the capital coefficients are not upward biased, and the returns to capital stated in Table 3 are reliable. Both the FE coefficients on labor and capital are very low, probably reflecting their downward bias due to weak identification and measurement error.

Table 7: Robustness: LP and FE estimator of the production function.

Method	Log labor Coefficient	S.E.	Log capital Coefficient	S.E.	N
LP	0.386***	(0.013)	0.109***	(0.013)	6610
FE	0.386***	(0.021)	0.042***	(0.010)	7857

Source: Authors' calculation based on the ENAHO firm panel.
Notes: LP estimates are obtained using the Stata procedure developed by Petrin, Poi, and Levinsohn (2004). LP standard errors (S.E.) are bootstrapped with 250 repl.; * p<0.10, ** p<0.05, *** p<0.01. + In all regressions, we drop influential outliers according to the procedure explained under Table 3.

In general, measurement error of both profits and capital tends to bias the estimated coefficients towards zero, so this source of bias does not give rise to major concerns in light of the significant strong effects. Moreover, we dropped influential outliers, as explained above. It is difficult to judge whether and how measurement error changes with higher or lower levels of profits and capital stock. We think there is little reason to assume that measurement error is less pronounced at lower levels of capital stocks, which would then partly explain higher returns at lower levels of capital.

Additional biases may arise from systematic firm exit. ME activities with low levels of capital stock are likely to be more vulnerable to shocks.[28] This bias should be less pronounced in our estimates, as we use an unbalanced panel. Nevertheless, we test for selection effects by first estimating the following equation for firm survival:

$$P(\text{Survival}_{it} = 1) = \text{F}(\theta k_{it} + \eta Z_{it} + \tau IS_{it}), \quad (17)$$

where Z_{it} is a vector of variables that proxy productivity (owner's age and its square, years of education, a dummy indicating male entrepreneurs). The term IS_{it} is a vector of variables that identify survival (see below). We estimate this equation as a simple probit. To correct for the selection problem, we then include the inverse of the mill's ratio in the augmented production function:

$$y_{it} = \alpha_i + \beta l_{it} + \gamma k_{it} + \rho Z_{it} + \frac{\phi\varphi(\hat{\theta} k_{it} + \hat{\eta} Z_{it} + \hat{\tau} IS_{it})}{\Phi(\hat{\theta} k_{it} + \hat{\eta} Z_{it} + \hat{\tau} IS_{it})} + u_{it}, \quad (18)$$

where ϕ denotes the standard normal pdf, and Φ is the standard normal cdf. The difficulty is to find variables that explain survival/exit, but are not correlated with value-added. We use a dummy indicating whether the entrepreneur is 60 years or older, and the number of small children in a household (less than one year old). That both variables may indeed be associated with involuntary business closure is supported by both instruments being jointly significant at the 10 percent level (not shown).[29] In the production function estimates the

[28] If MEs with a high capital stock are more likely to tolerate productivity shocks and remain in business, the capital coefficient would be biased downwards. Firms with a low capital stock that experience an adverse shock may move out of business while those with a higher capital stock react by reducing capital. This would imply an upward bias in the growth rate of small firms as only surviving firms are considered.

[29] Without IS_{it}, identification relies on the nonlinearity of the inverse mill's ratio. We have used a number of alternative instruments (including the number of household members, number of wage earners, transfer

inverse of the mill's ratios are not significant. Although the capital coefficient is indeed slightly higher once we control for selection, the difference is not significant (not shown). In contrast, the inverse of the mill's ratios are significant in the growth equation. Once we control for selection, small firms still grow considerably faster; however, the coefficient is slightly smaller, but not significantly different from the results presented above (not shown).

Another potential concern may be the validity of our risk measure that we construct from a combination of imperfect proxies (variation of sales, involuntary exit rate, idiosyncratic risk, and the latter aggregated at the sector level). As robustness check we therefore included each proxy separately and all proxies jointly in the estimations. In equations (14) and (15) all proxies – when included one by one – enter with the expected negative sign and are highly significant (Table A3 in the appendix reports the results for RE based idiosyncratic risk as the results for GMM based risk are very close). Once they are jointly included, only the idiosyncratic risk measure becomes insignificant in equation (15). In the estimation of equation (14) one proxy, the sector-level version of the idiosyncratic risk proxy, changes the sign, but is insignificant. These tests give us quite some confidence that our risk measure indeed provides an adequate summary measure of business risks. With respect to the size of the effects of risk, the inclusion of six sector dummies into equations (14) and (15) provides an important additional robustness check, as most risk proxies are constructed at the sector level and may therefore also capture other sectoral effects.[30] Not surprisingly, the effects of risk are being reduced and may be considered a lower bound estimate (see Table A4 in the appendix). This estimate still implies that a one standard deviation increase in the risk measure leads to a reduction of start-up capital and the growth rate by 20 and 10 percentage points, respectively. The coefficient and its interaction with wealth remain highly significant, whereas the interaction with enterprise age becomes insignificant.[31]

Table A5 reports the baseline system GMM specification for the production function estimation suggested by Eberhardt and Helmers (2010) where profit, labor and capital are predetermined, and capital and labor each lagged by one period are used as instruments. It differs from the specification in Table 4 chosen as that with most validated instruments where only profit is predetermined, and capital and labor each lagged by one period are used as instruments. The coefficients on labor and capital in Table A5 correspond to their expected values and are sufficiently close to the RE estimates. The coefficient on labor ranges between 0.56 and 0.7; only for the firms in the "hotels/restaurants" sector the coefficient is with a value of roughly 1 much larger. The coefficient on capital is on average around 0.13, which is similar to the estimates reported above. `Common factor restrictions' are still not rejected at 5 %.The Sargan and Hansen tests verifying the validity of the

---

incomes, household composition variables, a loss of employment, severe illness, criminal act or natural disaster). However, all of these variables are either not significant in the selection equation, or highly significant in the production function regression (after including the inverse mill's ratio).

[30] Sectors are disaggregated as finely as possible which means that they are not equivalent to common industry definitions (see for example Table 3).

[31] Other coefficients are not affected by the inclusion of sectoral dummies.

instruments are rejected, and so better specifications should be considered. However, the results on capital accumulation presented in Table A6 are extremely close to what we obtained using the main GMM specification supporting the overall robustness of the system GMM estimates.

6. Conclusions

In this paper, we theoretically and empirically analyze returns to capital and the dynamics of capital accumulation of MEs. We use a unique panel dataset of MEs from Peru and, in contrast to previous studies, find a pronounced effect of risk on capital accumulation. Our study starts from the observation that, in Peru, MEs appear to exhibit very high marginal returns – on average about 19 percent monthly – at low levels of capital, but these marginal returns decrease rapidly at higher levels. Yet, in a range of up to 130 USD of capital stock marginal returns are well above 10 percent monthly. More than 57 percent of MEs can be found in this range and may hence be able to realize these returns if provided the necessary capital. This result is in line with finding from a number of other contexts and is robust, as shown by the estimation of a number of alternative specifications. We then examine why capital stocks remain low and why MEs do not accumulate, for example by re-investing those high returns. Our analysis of the determinants of start-up capital and subsequent capital accumulation confirms that credit constraints explain a major part of the variation in firm growth. We find a very large effect of household non-business wealth on capital stocks of MEs. An increase in the wealth level by one standard deviation raises the growth rate of MEs by 35 percentage points in the GMM estimation. While the positive effect of wealth is a standard finding, it provides an interesting benchmark for the effects of risk that the present study – in our view – can show more convincingly than previous studies. We find risk to lead to considerably lower capital stocks and a slow process of capital accumulation. The effect is again sizable with a lower bound estimate that corresponds to 40 percentage points reduction in the growth rate for a one standard deviation increase in our preferred risk proxy. That MEs' investment decisions are heavily influenced by considering risks, is supported by pronounced interactions between wealth and risk. The presented evidence is consistent with poorly endowed entrepreneurs who operate in imperfect capital markets and a very risky environment. Hence, these entrepreneurs forego profitable investments, as they have to withhold liquidity and/or have very high discount rates. The "pure" risk effects may be efficient individual responses to prevailing business risks when insurance markets fail. However, the significant interaction between wealth and risk show that capital market imperfections reinforce the negative effect of risk on accumulation. This is an indication of important inefficiencies and unexploited potential in MEs caused by risk and credit constraints. However, this paper has little to say on how precisely this mechanism works and more research seems to be warranted that addresses these behavioral questions.

In general, the findings once again illustrate the great heterogeneity of informal activities (Cunningham and Maloney, 2001). While these activities may appear residual activities pursued for subsistence at first sight, the dynamics of an important part – if not the majority of the entrepreneurs in the Peruvian context – can be better described by theories of the firm; yet, of firms that operate under severe constraints. From a policy perspective, these results imply that credit constraints and risk leave the potential of many small-scale entrepreneurs unexploited. How these constraints can be overcome is difficult to say on the basis of our results. Of course, the strong effects of wealth provide a rationale for microcredit, but it is beyond the scope of this paper to judge whether efficient modes can be found to allocate capital to MEs in these small quantities. Access to simple savings accounts with light credit lines might be a possible solution. Savings accounts may also mitigate the effects of risk that we have identified. Whether such a measure will be effective in reducing the adverse effects of risk, however, depends on the precise nature of risks. In this regard, it is worth stressing that some risk is inherent to any business activity. This is also why more research is needed into how – possibly excessive – risk leads to sub-optimal investment decisions. Moreover, there is scope to analyze how household-related risks, for example health shocks, affect investment decisions. If these risks mattered, providing corresponding insurance, for example health and life insurance, may enhance capital accumulation in MEs. Both the determinants of ME investment behavior and the effectiveness and efficiency of specific policies are hence pertinent research questions since small-scale activities are likely to remain the main income source of the world's poor in decades to come.

References

Ahn, Taehyun. 2010. “Attitudes Toward Risk and Self-employment of Young Workers.” *Labour Economics*. 17(2): 434-442.

Akoten, John E., Yasuyuki Sawada, and Keijiro Otsuka. 2006. “The Determinants of Credit Access and Its Impacts on Micro and Small Enterprises: The Case of Garment Producers in Kenya.” *Economic Development and Cultural Change*, 54(4): 927-44.

Banerjee, Abhijit V., and Esther Duflo. 2014. “Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program.“ *Review of Economic Studies.* 81 (2): 572-607.

Banerjee, Abhijit V., and Kaivan Munshi. 2004. “How Efficiently is Capital Allocated? Evidence from the Knitted Garment Industry in Tirupur. ”*Review of Economic Studies*. 71(1): 19-42.

Belsley, David A., Edwin Kuh, and Roy E. Welsch. 1980. *Regression Diagnostics*. New York: Wiley.

Bigsten, Arne, Paul Collier, Stefan Dercon, Marcel Fafchamps, Bernard Gauthier, Jan Willem Gunning, Abena Oduro, Remco Oostendorp, Cathy Patillo, Måns Söderbom, Francis Teal, and Albert Zeufack.2003. “Credit Constraints in Manufacturing Enterprises in Africa.” *Journal of African Economies*. 12(1): 104-125.

Bloom, Nick, Stephen Bond and Richard van Reenen. 2007. Uncertainty and Investment Dynamics. *Review of Economic Studies*. 74: 391-415.

Blundell, Richard and Stephen Bond. 2000. “GMM Estimation with Persistent Panel Data: an Application to Production Functions.” *Econometric Reviews*. 19(3): 321-340.

Cabral, Luis .M.B., and José Mata. 2003. “On the Evolution of the Firm Size Distribution: Facts and Theory.” *American Economic Review* 93(4): 1075-1091.

Coad, Alex, and Jaganaddha Pawan Tamvada. 2012. “Firm growth and barriers to growth among small firms in India.” *Small Business Economics*. 39(2):383-400.

Cunningham, Wendy V., and William F. Maloney. 2001. “Heterogeneity among Mexico’s Microenterprises: An Application of Factor and Cluster Analysis.” *Economic Development and Cultural Change* (5): 131–156.

De Mel, Suresh, David McKenzie, and Christopher Woodruff. 2008. “Returns to Capital in Microenterprises: Evidence from a Field Experiment.” *The Quarterly Journal of Economics*. 123(4): 1329-1372.

De Mel, Suresh, David McKenzie, and Christopher Woodruff. 2012. “One-Time Transfers of Cash or Capital Have Long-Lasting Effects on Microenterprises in Sri Lanka.” *Science*. 335: 962-966.

Dixit, Avinash K., and Robert S. Pindyck. 1994. *Investment under Uncertainty*. Princeton University Press, Princeton.

Dromel, Nicolas L., Elie Kolakez, and Etienne Lehmann. 2010. "Credit Constraints and the Persistence of Unemployment." *Labour Economics*. 17(5): 823-834.

Dunn, Elizabeth, and Nicholas Kalaitzandonakes and Corinne Valdivia. 1996. "Risk and the Impacts of Microenterprise Services." Report submitted to USAID.

Evans, David S., and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy.* 97(4): 808-27.

Eberhardt, Markus and Christian Helmers. 2010. "Untested Assumptions and Data Slicing: A Critical Review of Firm-Level Production Function Estimators." Discussion paper 513. Oxford University.

Fafchamps, Marcel. 1999. "Rural Poverty, Risk, and Development." Report Submitted to the Food and Agricultural Organization.

Fafchamps, Marcel, David McKenzie, Simon Quinn, and Christopher Woodruff. 2011. "When is Capital Enough to Get Female Enterprises Growing? Evidence from a Randomized Experiment in Ghana." Policy Research Working Paper Series 5706, The World Bank.

Fafchamps, Marcel, and John L. Pender. 1997. "Precautionary Saving, Credit Constraints, and Irreversible Investment: Theory and Evidence from Semiarid India." *Journal of Business & Economic Studies,* 15(2): 180-194.

Fajnzylber, Pablo, William Maloney, and Gabriel Montes-Rojas. 2006. "Microenterprise Dynamics in Developing Countries: How Similar are They to Those in the Industrialized World? Evidence from Mexico." *World Bank Economic Review*. 20(3): 389-419.

Grimm, Michael, Jens Krüger, and Jann Lay. 2011. "Barriers to Entry and Returns to Capital in Informal Activities: Evidence from Sub-Saharan Africa." *Review of Income and Wealth*. 57(1): 27-53.

Janes, Laurin. 2013. Can Capital Grants Help Microenterprises Reach the Productivity Level of SMEs? Evidence from an Experiment in Sri Lanka. CSAE Working Paper WPS/2013-18. University of Oxford.

Jovanovic, Bojan. 1982. Selection and the Evolution of Industry. *Econometrica.* 50(3): 649-670.

Klapper, Leora, and Christine Richmond. 2011. "Patterns of Business Creation, Survival and Growth: Evidence from Africa." *Labour Economics*. 18(1): S32-S44.

Kremer, Michael, Jean N. Lee, and Jonathan Robinson. 2010. "The Return to Capital for Small Retailers in Kenya: Evidence from Inventories." Mimeo, Harvard University.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies*. 70(2): 317-341.

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

McKenzie, David, and Christopher Woodruff. 2008. “Experimental Evidence on Returns to Capital and Access to Finance in Mexico.” *World Bank Economic Review*. 22(3): 457-482.

McPherson, Michael A. 1996. “Growth of micro and small enterprises in southern Africa.” *Journal of Development Economics.* 48(21): 31–54.

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

Michelacci, Claudio, and Fabiano Schivardi. 2013. “Does Idiosyncratic Business Risk Matter for Growth?” Journal of the European Economic Association. 11(2): 343-368.

Nickell, Stephen. 1981. “Biases in Dynamic Models with Fixed Effects.” *Econometrica*. 49(6): 1417-1426.

Pender, John L.1996. “Discount rates and credit markets: Theory and evidence from rural India.” *Journal of Development Economics*. 50: 257-296.

Petrin, Amil, Brian P. Poi, and James Levinsohn. 2004. “Production function estimation in Stata using inputs to control for unobservables.” *Stata Journal*. 4(2): 113-123.

Ritter, Jay R. 1984. “The "Hot Issue" Market of 1980.”*The Journal of Business*. 57(2): 215-240.

Tybout, James R. 1983. “Credit Rationing and Investment Behavior in a Developing Country.” *Review of Economics and Statistics*. 65: 598-607.

Udry, Christopher, and Santosh Anagol. 2006. “The Return to Capital in Ghana.” *American Economic Review*. 96(2): 388-393.

Westbrook, M. Daniel, and James R. Tybout. 1993. “Estimating Returns to Scale with Large, Imperfect Panels: An Application to Chilean Manufacturing Industries.” *The World Bank Economic Review*. 7(1): 85-112.

Windmeijer, Frank. 2005. “A Finite Sample Correction for the Variance of Linear Efficient GMM estimators”. *Journal of Econometrics.* 126(1): 25-51.

Yesuf, Mahmud, and Randy Bluffstone. 2008.“Wealth and Time Preference in Rural Ethiopia.” Environment for Development Discussion Paper Series EFD DP 08-16.

Appendix

Proof of Proposition 2.

To find a closed form solution it is naturally to suppose that C_t and K_{t+1} will be proportional to the product $e^{\theta_t}K_t^{\alpha}$. The purpose then is to find π_C and π_K in two expressions:

$$C_t = (\pi_C e^{\theta_t} K_t^{\alpha})^{\frac{1}{\gamma}} \qquad \text{(A1)}$$

$$K_{t+1} = \pi_K e^{\theta_t} K_t^{\alpha} \qquad \text{(A2)}$$

Consider first the case without credit and risk constraints. Substitute (A1) into Euler condition (3):

$$(C_t)^{-\gamma} = \beta E_t (C_{t+1})^{-\gamma} \alpha e^{\theta_{t+1}} K_{t+1}^{\alpha-1},$$

$$\left(\pi_C e^{\theta_t} K_t^{\alpha}\right)^{-1} = \beta E_t \left(\pi_C e^{\theta_{t+1}} K_{t+1}^{\alpha}\right)^{-1} \alpha e^{\theta_{t+1}} K_{t+1}^{\alpha-1},$$

$$\left(\pi_C e^{\theta_t} K_t^{\alpha}\right)^{-1} = \alpha\beta E_t \left(\pi_C K_{t+1}\right)^{-1}.$$

Substituting (A2), we get

$$\left(\pi_C e^{\theta_t} K_t^{\alpha}\right)^{-1} = \alpha\beta \left(\pi_C \pi_K e^{\theta_t} K_t^{\alpha}\right)^{-1},$$

which implies that $\pi_K = \alpha\beta$.

We can then substitute (A1) and (A2) into the budget constraint to get

$$\left(\pi_C e^{\theta_t} K_t^{\alpha}\right)^{\frac{1}{\gamma}} = e^{\theta_t} K_t^{\alpha} - \alpha\beta e^{\theta_t} K_t^{\alpha} + \hat{A}_t$$

$$\pi_C = \frac{\left(e^{\theta_t} K_t^{\alpha} - \alpha\beta e^{\theta_t} K_t^{\alpha} + \hat{A}_t\right)^{\gamma}}{e^{\theta_t} K_t^{\alpha}}$$

That yields $C_t = (1-\alpha\beta) e^{\theta_t} K_t^{\alpha} + \hat{A}_t$.

In case with credit and risk constraints we substitute (A1) and (A2) into Euler condition (8):

$$\left(\pi_C e^{\theta_t} K_t^{\alpha}\right)^{-1} = \beta E_t \left(\pi_C e^{\theta_{t+1}} K_{t+1}^{\alpha}\right)^{-1} \alpha e^{\theta_{t+1}} K_{t+1}^{\alpha-1} - \alpha\beta e^{\theta_{t+1}} \frac{1}{K_{t+1}^{\alpha+1}} \cdot cov(u'(C_{t+1}), \varepsilon) - \lambda$$

Because $cov(u'(C_{t+1}), \varepsilon) = E[u'(C_{t+1}) \cdot \varepsilon] - E[u'(C_{t+1})] \cdot E[\varepsilon] = E[u'(C_{t+1}) \cdot \varepsilon] = E\left(\frac{\varepsilon}{\pi_C e^{\theta_{t+1}} K_{t+1}^{\alpha}}\right)$,

the Euler condition can be rewritten as

$$\left(\pi_C e^{\theta_t} K_t^{\alpha}\right)^{-1} = \alpha\beta E_t \left(\pi_C K_{t+1}\right)^{-1} - \alpha\beta e^{\theta_{t+1}} \frac{1}{K_{t+1}^{\alpha+1}} \cdot E\left(\frac{\varepsilon}{\pi_C e^{\theta_{t+1}} K_{t+1}^{\alpha}}\right) - \lambda,$$

$$\frac{1}{\pi_C e^{\theta_t} K_t^{\alpha}} = \frac{\alpha\beta}{\pi_C \pi_K e^{\theta_t} K_t^{\alpha}} - \alpha\beta E\left(\frac{\varepsilon}{\pi_C (\pi_K e^{\theta_t} K_t^{\alpha})^{2\alpha+1}}\right) - \lambda.$$

As $E(\varepsilon) = 0$, then $\alpha\beta E\left(\frac{\varepsilon}{\pi_C (\pi_K e^{\theta_t} K_t^{\alpha})^{2\alpha+1}}\right) = 0$, and we can isolate π_C:

$$\pi_C = \frac{\alpha\beta}{\lambda\pi_K e^{\theta_t} K_t^{\alpha}} - \frac{\pi_K}{\lambda\pi_K e^{\theta_t} K_t^{\alpha}} \qquad \text{(A3)}$$

Substituting (A3) into the budget constraint we obtain

$$\left(\frac{\alpha\beta}{\pi_K} - 1\right)^{\frac{1}{\gamma}} \left(\frac{1}{\lambda}\right)^{\frac{1}{\gamma}} = e^{\theta_t} K_t^{\alpha} + \varepsilon \frac{e^{\theta_t}}{K_t^{\alpha}} - \pi_K e^{\theta_t} K_t^{\alpha} + \hat{A}_t$$

As $K_{t+1} = \pi_K e^{\theta_t} K_t^{\alpha}$ we can rewrite this expression to obtain (10)

$$K_{t+1} = \pi_K e^{\theta_t} K_t^{\alpha} = e^{\theta_t} K_t^{\alpha} + \varepsilon \frac{e^{\theta_t}}{K_t^{\alpha}} + \hat{A}_t - \left(\frac{\alpha\beta}{\pi_K} - 1\right)^{\frac{1}{\gamma}} \left(\frac{1}{\lambda}\right)^{\frac{1}{\gamma}}.$$

Similarly, we can isolate π_K from the Euler condition:

$$\pi_K = \alpha\beta - \lambda\pi_K\pi_C e^{\theta_t} K_t^{\alpha} = \alpha\beta - \lambda\pi_K C_t^{\gamma},$$

$$\pi_K = \frac{\alpha\beta}{1 + \lambda C_t^{\gamma}}.$$

Multiplying by $e^{\theta_t} K_t^{\alpha}$, we get

$$K_{t+1}^{withconstr} = \frac{K_{t+1}^{w/outconstr}}{1 + \lambda C_t^{\gamma}},$$

$$\frac{K_{t+1}^{w/outconstr}}{K_{t+1}^{withconstr}} = 1 + \lambda C_t^{\gamma}.$$

Proof of Corollary 1.

To make explicit how capital accumulation depends on productivity, wealth, risk and credit constraints rearrange (10) in the following way:

$$K_{t+1} = e^{\theta_t} K_t^{\alpha} + \varepsilon \frac{e^{\theta_t}}{K_t^{\alpha}} + \hat{A}_t - \left(\frac{\alpha\beta e^{\theta_t} K_t^{\alpha}}{K_{t+1}} - 1\right)^{\frac{1}{\gamma}} \left(\frac{1}{\lambda}\right)^{\frac{1}{\gamma}} \qquad \text{(A4)}$$

Without credit constraints and risk $K_{t+1} = \alpha\beta e^{\theta_t} K_t^{\alpha}$, so that (A4) can be rewritten as:

$$K_{t+1}^{withconstr} = e^{\theta_t} K_t^{\alpha} + \varepsilon \frac{e^{\theta_t}}{K_t^{\alpha}} + \hat{A}_t - \left(\frac{K_{t+1}^{w/outconstr}}{K_{t+1}^{withconstr}} - 1\right)^{\frac{1}{\gamma}} \left(\frac{1}{\lambda}\right)^{\frac{1}{\gamma}}. \qquad \text{(A5)}$$

Risk aversion γ is supposed to be any numeric value so $\frac{K_{t+1}^{w/outconstr}}{K_{t+1}^{withconstr}} - 1$ is positive to assure the root operation. That leads to $\frac{K_{t+1}^{w/outconstr}}{K_{t+1}^{withconstr}} > 1$ or $K_{t+1}^{w/outconstr} > K_{t+1}^{withconstr}$. Although the proofs are based on special cases of utility and production functions, they are done without loss of generality.

Table A1: Structure of the panel

Year	Hh. visited	Hh. not interviewed	Hh. observed in previous period	Hh. interviewed
2002	6257	847	.	5410
2003	4217	688	3068	3529
2004	6490	1141	2787	5349
2005	6778	1469	4146	5309
2006	6593	1182	4496	5411

Source: Authors' calculation based on the ENAHO household panel.

Table A2: Risk proxies, correlation matrix

	Sales variation+++	exit rate+++	RES+++	RES	**Component 1**	**Component 1**
Sales variation+++	1.000				**0.548**	**0.540**
exit rate+++	0.290***	1.000			**0.595**	**0.552**
RE RES+++	0.311***	0.330***	1.000		**0.588**	**0.590**
RE RES	0.051***	0.054***	0.165***	1.000	**-**	**0.235**

Source: Authors' calculation based on the ENAHO firm panel.
Notes: * p<0.10, ** p<0.05, *** p<0.01. +++ Variable is measured at the sector level. The last columns on the right show the eigenvector associated to the first component (without and with the idiosyncratic risk).

Table A3: Risk proxies, regressions

Dep. var.:	start-up *k*				Δ k				
Sales var.°	-0.549***			-0.490***	-0.182***				-0.129***
	(0.056)			(0.058)	(0.030)				(0.030)
exit rate°		-7.589***		-6.364***		-2.762***			-2.134***
		(0.874)		(0.878)		(0.412)			(0.445)
RE RES°			-2.549**	1.238			-3.047***		-1.651***
			(1.062)	(1.087)			(0.513)		(0.556)
RE RES								-0.194**	-0.113
								(0.082)	(0.083)
R2	0.228	0.214	0.173	0.253	0.178	0.179	0.173	0.170	0.198
N	1171	1173	1174	1173	2446	2447	2450	2436	2435

Source: Authors' calculation based on the ENAHO firm panel.
Notes: Regressions include all additional explanatory variables which are contained in the main regressions of equation (6) and (7). In all regressions, we drop influential outliers according to the procedure explained in the note to Table 3. +++ Variable is measured at the sector level. Robust standard errors in parenthesis * p<0.10, ** p<0.05, *** p<0.01.

Table A4: Estimated functions of start-up capital and accumulation with industry dummies (based on RE production function estimates from Table 3)

Dependent variable:	*start-up k (OLS)*	*Δ k (RE)*	
Log capital		-0.553***	-0.552***
		(0.023)	(0.030)
Wealth index	0.201***	0.159***	0.184***
	(0.032)	(0.015)	(0.022)
Risk index	-0.199***	-0.139***	-0.137***
	(0.058)	(0.032)	(0.042)
Wealth * Risk	0.037*	0.027**	0.028**
	(0.020)	(0.012)	(0.008)
Enterprise age		-0.023***	-0.023**
		(0.009)	(0.010)
Enterprise age^2		0.0005**	0.0005*
		(0.00024)	(0.00028)
TFP		0.245***	0.247***
		(0.067)	(0.062)
Enterprise age * risk			-0.00057
			(0.003)
Enterprise age * wealth			-0.003**
			(0.001)
Year dummies	Yes	Yes	Yes
Outlier controlled+	Yes	Yes	Yes
R2	0.388	0.238	0.241
N	1144	2443	2441

Source: Authors' calculation based on the ENAHO firm panel.
Notes: Regressions include all variables stated under Table 5. In all regressions, we drop influential outliers according to the procedure explained in a note to Table 3. Sector dummies represent all sectors as in Table 3 with petty trading as baseline. Robust standard errors are bootstrapped and presented in parentheses * p<0.10, ** p<0.05, *** p<0.01

Table A5: Production function estimates for Peruvian industries (GMM baseline specification)

Sectors	All	M.& food	Constr.	W./R. shops	Petty trad.	H.& R.	Tran.	Oth. serv.
Log labor	0.682***	0.701***	0.571***	0.573***	0.665***	1.052***	0.557***	0.614***
SE	(0.043)	(0.094)	(0.113)	(0.130)	(0.069)	(0.120)	(0.113)	(0.143)
Log capital	0.134***	0.242***	0.069	0.047	0.091***	0.021	0.183***	0.142***
SE	(0.013)	(0.033)	(0.052)	(0.068)	(0.028)	(0.046)	(0.047)	(0.041)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Outlier[+]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Average K	106	175	29	124	50	76	1181	99
Average Y	122	112	161	239	99	139	159	118
Average Y/K	8.8	4.2	24.9	11.5	12.6	6.7	0.5	4.4
M. returns	0.154	0.155	0.376	0.091	0.182	0.039	0.025	0.169
N	2244	284	114	78	943	310	345	169
N of groups	856	132	50	40	404	141	142	83
Sargan P-value	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000
Hansen P-value	0.000	0.094	0.304	0.605	0.000	0.090	0.005	0.181
AR(1)	0.057	0.662	0.486	0.303	0.020	0.252	0.114	0.241
AR(2)	0.106	0.582	0.810	0.181	0.745	0.931	0.845	0.137
COMFAC	0.147	0.225	0.549	0.964	0.835	0.183	0.051	0.478

Source: Authors' calculation based on the ENAHO firm panel.

Notes: In all regressions, we drop influential outliers from our sample (and sub-samples) that we identify by the DFITS-statistic obtained from a precedent OLS regression (See Belsley, Kuh, and Welsch, 1980). We use a cut-off-value $|DFITS| = 3\sqrt{\mathrm{k/N}}$, with$k$, the degrees of freedom (plus 1) and N, the number of observations. This reduces the sample by 1-2 percent. Marginal returns to capital $\frac{dY}{dK} = \gamma * \frac{Y}{K}$ are evaluated at Y=exp(mean(log Y)) and K=exp(mean(log K)). The estimator employed is BB – Blundell and Bond (1998) System GMM with predetermined profit, capital and labor, and lagged capital and labor as instruments. We use the 2-step estimator with the Windmeijer (2005) correction. The sample includes all entrepreneurs observed during at least three time periods. COMFAC is a minimum distance test of the common factor restrictions. P-values reported for all test statistics. For the BB estimator the Difference-in-Sargan and Hansen statistics test the validity of all the additional instruments employed for the levels equation. All long-run coefficients are computed using the Minimum Distance estimator (code from Mans Söderbom) and employing the Delta method to obtain standard errors (absolute values reported).Robust standard errors in parentheses (SE)* p<0.10, ** p<0.05, *** p<0.01.

Table A6: Estimated functions of start-up capital and accumulation (based on GMM production function estimates from Table A5)

Dependent variable:	*start-up k (OLS)*	*Δ k (GMM)*	
Log capital		-0.844***	-0.855***
		(0.083)	(0.083)
Wealth index	0.126*	0.206***	0.220***
	(0.068)	(0.029)	(0.036)
Risk index	-0.352**	-0.425***	-0.503***
	(0.139)	(0.055)	(0.072)
Wealth * Risk	0.177***	0.0413**	0.0408**
	(0.059)	(0.016)	(0.016)
Enterprise age		-0.0448***	-0.0450***
		(0.014)	(0.014)
Enterprise age^2		0.00076*	0.00077*
		(0.0004)	(0.0004)
TFP		0.223***	0.220***
		(0.0495)	(0.049)
Enterprise age * risk			0.008*
			(0.004)
Enterprise age * wealth			-0.0008
			(0.0024)
Year dummies	Yes	Yes	Yes
Outlier controlled+	Yes	Yes	Yes
N	136	1217	1208
N of groups		747	743
R2	0.241		
Sargan P-value		0.252	0.348
Hansen P-value		0.602	0.642

Source: Authors' calculation based on the ENAHO firm panel.
Notes: VA and capital are predetermined. The regression results are extremely close to those reported in Table 6.

Figure A1: Descriptives