Testing for Credit Constraints in Entrepreneurship
Sam Schulhofer-Wohl

For the Applications workshop on February 7, I will present primarily material from sections 3 and 4 (pages 12 to 25) of the attached paper.

Testing for Credit Constraints in Entrepreneurship

Sam Schulhofer-Wohl*
September 3, 2004

Abstract

Common tests for credit constraints are highly sensitive to econometric assumptions. Tests of the relationship between wealth and entry into entrepreneurship require strong assumptions about the distribution of unobserved ability, and the results change when the definition of entrepreneurship changes. Tests of the relationship between entrepreneurs' productivity and wealth can produce spurious evidence of credit constraints if heterogeneity in technology is ignored. Using data from Nicaragua, I illustrate these problems, show how to make the productivity test more robust, and demonstrate a semiparametric method for measuring the implicit interest rates faced by entrepreneurs at different wealth levels. As many as 80% of Nicaraguan entrepreneurs face financial constraints. The implicit interest rate for high-wealth entrepreneurs is at least 35% below the rate at the median wealth level.

*Department of Economics, University of Chicago, sschulh1@uchicago.edu. I thank Robert Townsend and Sergio Urzua for helpful comments.

1 Introduction

An economy cannot achieve the highest possible growth rate unless resources – capital, labor and so forth – are used in the activities where they are most productive. An inefficient allocation of resources "leaves money on the table": Relatively unproductive activities waste inputs, while highly productive opportunities lie abandoned. Because small businesses and farms make up a large fraction of most economies, major theoretical and empirical literatures have investigated whether markets efficiently allocate resources to small enterprises. The allocation of capital is of particular interest and importance. Without smoothly functioning credit markets, the only capital available to an entrepreneur is her own wealth. Changes in the distribution of wealth can therefore dramatically affect whether capital is used in the most productive activities and, thus, how fast the economy grows.[1]

One commonly used test for an efficient allocation of capital starts from the observation that if access to credit is imperfect, poor people will be less likely to start businesses because they cannot provide the necessary capital themselves. Thus, a researcher can estimate the probability that a person starts a business as a function of her wealth. If the probability is increasing in wealth, then, as in Evans and Jovanovic (1989), the results might be interpreted as evidence of credit constraints.

Such a test presumes that people's skill at entrepreneurship is unrelated to wealth. Otherwise, the probability of starting a business might increase with wealth simply because wealthy people are more talented entrepreneurs. This theoretical problem accompanies a practical concern: Who counts as an entrepreneur? For example, what fraction of a rural family's crops must be sold for cash instead of eaten at home before the family is engaged in an agricultural business instead of subsistence farming? This paper demonstrates that, even setting aside concerns about the relationship between wealth and entrepreneurial talent, small changes in the definition of entrepreneurship can greatly change conclusions about credit constraints.

A more direct test for efficient credit markets can avoid these problems. A well-known approach in the corporate finance literature uses the fact that an efficient credit market equalizes the marginal

[1]Well-known models of the relationship between wealth distribution and economic growth include Aghion and Bolton (1997), Banerjee and Newman (1993), Galor and Zeira (1993), and Lloyd-Ellis and Bernhardt (2000).

product of capital across all firms. Thus, if firms with lower cash flow or less wealthy owners have higher marginal products of capital, the credit market must be inefficient.[2] Under some assumptions, a firm's marginal product of capital is proportional to the ratio of the firm's output to its capital stock. Negative correlations between firms' output-to-capital ratios and owners' wealth may therefore signal credit constraints.

Paulson and Townsend (2004) find that output-to-capital ratios decrease with owners' wealth among household businesses in Thailand, and interpret this as evidence of credit constraints. The interpretation relies on the assumption that all firms' technologies are equally capital-intensive. I present evidence that this assumption is unlikely to hold and that omitting controls for variation in technology across industries can create the appearance of credit constraints where none exist. Estimators that control for the industry in which a firm operates can produce more reliable evidence on whether entrepreneurs face credit constraints.

A statistical rejection of efficient capital markets does not, in itself, indicate the magnitude of the credit constraints. However, each entrepreneur's marginal product of capital corresponds to a shadow interest rate. In the absence of credit constraints, the shadow interest rate is the same as the market interest rate. In the presence of credit constraints, the shadow interest rate is higher than the market rate; it is the rate that would make the entrepreneur's constrained investment choice rational in a market with no constraints. A survey cannot measure the shadow interest rate directly by asking entrepreneurs about the interest rates they pay, because credit constraints may make loans unavailable at any rate. Instead, the shadow interest rate must be recovered from entrepreneurs' investment behavior. I show how to estimate shadow interest rates as a function of wealth, using data on output-to-capital ratios and the partially linear estimator of Robinson (1988). Differences in shadow interest rates across wealth levels indicate the degree to which capital is allocated inefficiently.

Banerjee and Duflo (2002) similarly estimate shadow interest rates for firms in a developing country. My method differs from theirs in two ways: I allow firms' technology to differ, and I estimate a schedule of shadow interest rates for the entire wealth distribution rather than a single

[2] Fazzari, Hubbard and Petersen (1988) is a seminal paper in this area; see Kaplan and Zingales (1997) for a critical response.

shadow interest rate corresponding to a particular natural experiment that changes the availability of credit in a specific way. The price of this greater generality is that I measure shadow interest rates *given* the joint distribution of wealth, productivity and technology and cannot say how a given entrepreneur's shadow interest rate would change if her wealth changed, holding all else fixed.

I illustrate tests for credit constraints with panel data from the 1998 and 2001 Living Standards Measurement Surveys in Nicaragua. Because the Nicaragua dataset contains detailed information on households' access to the financial system, I can examine whether tests for credit market imperfections correctly identify as unconstrained the households that report facing no financial constraints. The Nicaragua surveys also include detailed codes for entrepreneurs' industries, which allow me to examine the importance of technological variation. To preview the results: The data show that as many as 80% of Nicaraguan entrepreneurs face financial constraints and that the shadow interest rate for high-wealth entrepreneurs is at least 35% below the shadow interest rate at the median wealth level.

The paper proceeds as follows: Section 2 examines theoretical and practical problems with tests for credit constraints that examine the probability that a person starts a business. Section 3 shows how, in theory, entrepreneurs' output-to-capital ratios can be used to test for credit constraints. Section 4 applies various versions of this test, measures the degree of credit constraints in Nicaragua, and investigates the credibility of the test by applying it to data on entrepreneurs who are known to be unconstrained. Section 5 concludes. The appendix describes the data in detail.

2 Occupation choice tests

Credit markets can be incomplete in many ways. People may have no access to credit; they may be permitted to borrow only a limited amount because lenders fear they will not repay; or their borrowing may be constrained by the concern that people who invest little of their own money have no incentive to work hard. Regardless of the reason, however, credit constraints imply that someone who must borrow to start a business may have lower profits than someone who is wealthy enough to self-finance, because the borrower may have to employ less than the optimal amount of capital. By contrast, if credit markets are complete, both the borrower and the person who self-finances will employ the same quantity of capital and have the same profits.

Credit constraints thus mean that if a poor person and a wealthy person have equal skills or talents as entrepreneurs, the wealthy person may be more likely to start a business. In practice, however, a person's potential skill as an entrepreneur may be unobservable. This leads to a question of identification: Given data only on people's wealth, their decision to start a business and perhaps some covariates such as their education – but not on their skill as entrepreneurs – can one say anything about the presence of credit constraints?

2.1 Identification

I consider the identification problem in a very general context. The model I use includes as special cases perfect credit markets, the static limited commitment model of Evans and Jovanovic (1989), the dynamic limited commitment model of Buera (2003a), the moral hazard model of Aghion and Bolton (1997), and the startup costs model of Lloyd-Ellis and Bernhardt (2000).

There are two occupations: wage work and entrepreneurship. A household chooses whichever occupation is more profitable. A vector of observable variables x can affect profits in the two occupations, as can an unobservable variable, θ, which I will call ability. In addition, I assume that there are data on each household's wealth, b, and on an indicator variable d whose value is 1 if a household chooses entrepreneurship and is 0 otherwise.

I assume that, conditional on ability, wealth cannot influence a household's productivity in either occupation except by affecting the quantity of capital that the household may employ in business. Let $f(k;\theta,x)$ be the output of a household with ability θ and characteristics x if it starts a business and employs k units of capital. Let $R(k)$ be the return to k units of capital that are not invested in a business, and let $w(\theta,x)$ be the wage of a household with ability θ and characteristics x if it chooses wage work. Then the profits that a household can attain in entrepreneurship are

$$g_e(\theta|x,b) \equiv \max_k \{f(k;\theta,x) + R(b-k)\}. \tag{1}$$

Similarly, the total earnings of a household that chooses wage work are $g_w(\theta|x,b) \equiv w(\theta,x) + R(b)$. The difference in profits between the two occupations is $g(\theta|x,b) \equiv g_e(\theta|x,b) - g_w(\theta|x,b)$, and the household chooses entrepreneurship if and only if $g(\theta|x,b) \geq 0$. I assume that $g(\theta|x,b)$ is a monotonically increasing function of θ for each pair (x,b). Therefore, g can be inverted to find a critical ability level $\theta^*(x,b)$ such that a household with characteristics x and wealth b chooses entrepreneurship if and only if it has ability $\theta \geq \theta^*(x,b)$.

This model is essentially the economy of Roy (1951). If g_e and g_w are one-period profit and earnings functions, then it is a static model. But it equally well describes the instantaneous choice of occupation in a dynamic model such as that of Buera (2003a), if g_1 and g_0 are viewed as the value functions of households in each sector.

If there are perfect credit markets, the household can borrow or lend arbitrary quantities of capital at a constant interest rate r. That is, $R(k) = rk$ for all k positive or negative. Then

$$g_e(\theta|x,b) = \max_k \{f(k;\theta,x) - rk\} + rb \tag{2}$$

and

$$g_w(\theta|x,b) = w(\theta,x) + rb. \tag{3}$$

The difference in profits between the two occupations is

$$g(\theta|x,b) = \max_k \{f(k;\theta,x) - rk\} - w(\theta,x), \tag{4}$$

which does not depend on wealth b. The critical ability level $\theta^*(x,b)$ therefore also does not depend on wealth when there are no credit constraints.

Credit constraints of any sort imply that $R(k)$ is a nonlinear and possibly nonfinite function. If the household can lend but not borrow, for instance, then $R(k) = -\infty$ for $k < 0$. Nonlinearities in $R(k)$ mean that $g(\theta|x,b)$ and $\theta^*(x,b)$ depend on b. The form of this relationship depends on the specific source of credit constraints.

A test for credit market imperfections using choice data thus amounts to a test for whether $g(\theta|x,b)$ and $\theta^*(x,b)$ depend on wealth. One can determine whether credit constraints exist only if one can determine the form of $g(\theta|x,b)$ or $\theta^*(x,b)$. A researcher who computes the probability of starting a business as a function of wealth, say by estimating a probit model or a nonparametric relationship, is implicitly trying to estimate $g(\theta|x,b)$ and $\theta^*(x,b)$.

Let the cumulative distribution of ability conditional on wealth and other observables be $H(\theta|x,b)$. Since the data include the occupation choice d as well as x and b, the probability $\Pr[d=1|x,b]$ that a household chooses entrepreneurship given characteristics x and wealth b is identified. By the definition of the critical ability level θ^*,

$$\Pr[d=1|x,b] = \Pr[\theta \geq \theta^*(x,b)|x,b] = 1 - H[\theta^*(x,b)|x,b]. \tag{5}$$

Thus, if H is invertible and its form is known, the critical ability level $\theta^*(x,b)$ can be recovered from the probability of choosing entrepreneurship:

$$\theta^*(x,b) = H^{-1}\big[1 - \Pr[d=1|x,b]\big]. \tag{6}$$

Conversely, without some assumptions about the distribution of ability H, it is impossible to identify the critical ability level $\theta^*(x,b)$. Any proposed function $\theta^*(x,b)$ can be made consistent with the data by choosing an appropriate distribution $H(\theta|x,b)$: Just choose H to satisfy equation (6) for each pair (x,b). Without assumptions about H, the model thus is nonparametrically unidentified.

A lengthy empirical literature attempts to solve the identification problem. One solution is to look for variation in wealth that is unrelated to ability. Bequests have been the most commonly used source of variation; see, for example, Blanchflower and Oswald (1998) and Holtz-Eakin, Joulfaian and Rosen (1994a, 1994b). But exogenous variation is hard to find. Bequests have been the most commonly used source of variation. However, Hurst and Lusardi (2004) show that entrepreneurial activity is correlated not only with past bequests but also with future bequests, suggesting that bequests may be related to ability.

Another solution is to study a group of people for whom wealth is assumed to be independent of ability. Matzkin (1992) shows that the occupation choice model is nonparametrically identified under this assumption. Late in the life cycle, ability is likely to be correlated with wealth either because those with high ability earn rents or because those with high ability save more money in order to be able to start businesses, as in the dynamic model of Buera (2003a). Young households, which have not had much time to accumulate wealth, thus are the leading candidate for a group where wealth and ability are independent; see Jeong and Townsend (2003). However, because young people have had little time to acquire wealth on their own, what wealth they do have may come substantially from their parents. If parents can transmit some entrepreneurial ability to children, either genetically or in the course of child-rearing, then wealth will be related to ability even among young households, and this identifying assumption will be incorrect.

A third method concedes that some relationship between talent and wealth is unavoidable and attempts to model this relationship. In this structural approach, the researcher investigates whether a particular parametric relationship between talent, wealth and other variables and particular forms of the profit functions g_1 and g_0 are sufficient to explain the observed relationship between entrepreneurship and wealth. Evans and Jovanovic (1989) assume that entrepreneurs' profits are log-linear in capital and that ability follows a log-normal distribution whose mean depends monotonically on wealth b. Paulson and Townsend (2003) allow the mean of ability also to depend on education.

Results using the structural approach depend crucially on the accuracy of the assumed parametric relationship. Incorrect distributional assumptions can attribute either too much or too little of the variation in entrepreneurship to variation in talent, rather than variation in wealth. This can lead to mistaken conclusions about credit constraints. Consider the following four structural models:

1. Ability is normally distributed with mean 0 and variance 1 regardless of wealth b or other characteristics x. The critical level of ability for starting a business is $\theta^*(x, b) = -x'\beta - \gamma b$ for some parameters β and γ.

2. Ability is normally distributed with mean $x'\beta$ and variance 1 regardless of wealth b. The critical level of ability for starting a business is $\theta^*(x, b) = -\gamma b$.

3. Ability is normally distributed with mean γb and variance 1 regardless of other characteristics x. The critical level of ability for starting a business is $\theta^*(x, b) = -x'\beta$.

4. Ability is normally distributed with mean $x'\beta + \gamma b$ and variance 1. The critical level of ability for starting a business is $\theta^*(x, b) = 0$.

In all four of these models, the probability of starting a business is

$$\begin{aligned} \Pr[d = 1|x, b] &\equiv \Pr[\theta \geq \theta^*(x, b)|x, b] \\ &= \Pr[\theta - \mathrm{E}(\theta|x, b) \geq \theta^*(x, b) - \mathrm{E}(\theta|x, b)|x, b] \\ &= \Pr[\underbrace{\theta - \mathrm{E}(\theta|x, b)}_{\sim N(0,1)} \geq -x'\beta - \gamma b|x, b] = \Phi(x'\beta + \gamma b), \end{aligned} \tag{7}$$

where Φ is the standard normal cumulative distribution function. Thus a researcher would estimate the parameters β and γ of any of these four models by estimating a probit equation relating occupation choice to characteristics x and wealth b,

$$\Pr[d = 1|x, b] = \Phi(x'\beta + \gamma b). \tag{8}$$

The only difference between the models is in how the resulting estimates of β and γ will be interpreted. According to either the first or the second model, if $\gamma > 0$, the critical ability level θ^* increases with wealth b and there are credit constraints. In the third and fourth models, θ^* does not depend on wealth, implying that there are no credit constraints, but if $\gamma > 0$ then mean ability rises with wealth. The identification problem is simply that the data cannot tell us which of the four models is correct. A positive coefficient on wealth in the probit equation does not prove that there are credit constraints unless one assumes *a priori* that the mean level of ability does not depend on wealth.

Additional data can ease the identification problem. Buera (2003b) shows that occupation choice models can be identified given data on the earnings of both entrepreneurs and wage workers, essentially because earnings data reveal the individual's ability. However, high-quality earnings data are not always available, especially for day laborers, whose income fluctuates frequently, and subsistence farmers, much of whose output is unpriced. The productivity-based tests in Sections 3 and 4 employ data on the earnings of entrepreneurs alone to attempt to identify credit constraints.

2.2 Defining an entrepreneur

The model above assumes that all people can be clearly classified as either entrepreneurs or not entrepreneurs. Researchers investigating entrepreneurship in developing countries have often viewed farming as a default or subsistence occupation, so that only non-farm enterprises count as businesses. The definition of a business must, however, be sensitive to the details of a country's economy. For example, Nicaragua, the country used as an illustration in this paper, has not only subsistence farmers but also large coffee and banana plantations. Counting the wealthy owners of these plantations as subsistence farmers may lead to mistaken conclusions about the effect of wealth on starting a business. I therefore compare two definitions of a business:

Definition 1. *A business is any non-farm enterprise. This is the definition used by Paulson and Townsend (2003, 2004) in Thailand and Tejerina (2004) in Nicaragua, among others.*

Definition 2. *A business is any non-farm enterprise that uses capital, as well as any farm that sells output for cash and employs physical capital, such as machinery or buildings. About 32% of farming households in Nicaragua in 1998 qualify as businesses under this definition.*

The second definition undoubtedly includes some farms that are not really businesses, just as the first definition undoubtedly excludes some farms that are businesses. Because it may be impossible to draw a perfect dividing line, a test for credit constraints should be robust to changes in the definition.

If, in a cross-section of people, the wealthy are more likely to own businesses, it is impossible to know whether wealth made it easier for them to start businesses or whether they grew wealthy precisely because they had already started businesses. To estimate the probability of starting a business as a function of wealth, wealth therefore must be measured before entrepreneurship. A researcher can either construct retrospective measures of wealth in cross-sectional data, as in Paulson and Townsend (2004), or else employ panel data, measuring wealth in the first wave of the panel and entrepreneurship in the second wave. Each of these methods has drawbacks. Retrospective wealth data are likely to be measured with error and may be contaminated by current wealth, while panel data may suffer from attrition. Here, I use panel data, measuring Nicaraguan households' wealth in 1998 and their business activity in 2001. The appendix describes the data in detail.

Among the 3,890 households for which there are complete data in 1998, only 2,932, or 75%, were interviewed again in 2001.[3] To account for this attrition, I reweight each household interviewed in 2001 by the inverse of the probability that a household with similar characteristics in 1998 would remain in the sample three years later. This procedure is adequate if households that were and were not interviewed in 2001 differ only in observable attributes, or if any unobservable differences are unrelated to starting a business. If any unobserved differences are related to starting a business, the results will be biased in an unknown direction. The appendix gives details of how I compute these weights. However, all of the results presented in the paper are insensitive to the exact method of computing the weights or to ignoring attrition altogether.

Table 1 summarizes the data. Changing how farms are classified is likely to be much more important in areas where there are many farms, so I divide the households according to whether they live in rural or urban areas. (Slightly fewer than half of Nicaraguan families live in rural areas.) I drop households whose net assets in 1998 were less than or equal to zero because these asset values are likely to result from measurement error. The summary statistics show the relationship between wealth and starting a business. For both definitions of a business, and in both urban and rural areas, households that started businesses between 1998 and 2001 had higher median wealth than households that did not start businesses.

The relationship can be seen more precisely by estimating the probability that a household owned a business in 2001, given that it did not own a business in 1998, as a function of wealth in 1998. I use nonparametric local linear regressions (Fan, 1992) to estimate this function using each of the two definitions of a business in turn. For a given value of wealth, say b^*, the local linear regression procedure estimates a weighted linear regression of an indicator for whether the household started a business, d_i, on the household's wealth, b_i. The weights in this regression are highest for observations with wealth b_i closest to b^*. As the sample size rises, the weights emphasize

[3] This attrition may be in part a consequence of Hurricane Mitch. The November 1998 storm, one of the worst ever to hit Central America, came shortly after the 1998 survey was completed and left 1 million Nicaraguans – about 20% of the population – homeless.

an increasingly small neighborhood around b^*.[4] Thus, in the limit, only observations with wealth equal to b^* would contribute to the estimated probability of starting a business at wealth b^*. I repeat this procedure on a grid of 500 values of b^* to estimate the probability of starting a business over the entire support of the wealth distribution.[5] I find confidence intervals by bootstrapping: I draw a large number of samples from the original data with replacement, compute the estimator for each of these new samples, and use the fifth to 95th percentiles of these estimates at each value of b^* as a 90% confidence interval for the true value at b^*.

The estimates for urban areas are insensitive to the definition, and I do not report them here. Figure 1 displays the estimates for rural areas. The top panel shows the estimated probability of starting a business as a function of wealth, under each definition. The bottom panel shows the estimated density of wealth among families that did not own a business in 1998. (The asymptotic variance of the local linear regression estimator is lowest where the density is highest.) The probability of starting a non-farm business increases only slightly with wealth, if at all. An analyst using Definition 1 thus would be unlikely to conclude that substantial credit constraints exist in rural Nicaragua. However, the probability of starting any business, including a farm business, slopes sharply upward. An analyst using Definition 2 would conclude that credit constraints have a large effect in rural Nicaragua.

[4] Specifically, the weights used here are

$$w_i = K\left(\frac{b_i - b^*}{h}\right),$$

where $K(u) = 0.75(1-u^2)$ for $u \in (-1, 1)$ and $K(u) = 0$ otherwise, the Epanechnikov kernel. (Because the data come from a stratified sample, I then multiply w_i by the sampling weights.) The number h is called the bandwidth. The smaller the bandwidth, the smaller the neighborhood of observations that affect the estimate at b^*. The bandwidth must be chosen to balance bias, which increases with h, against variance, which decreases with h. This can be done by a variety of data-driven methods or by visually inspecting the results from using various values of h. I use the latter "eyeball" method; the chosen bandwidths are reported with each estimate.

[5] Because the wealth distribution is heavily skewed, I take the log of wealth before computing the estimates. This does not affect the asymptotic properties of the estimator, but a more symmetric distribution of the regressor improves the finite sample properties when the same bandwidth is used for all values of b^*.

3 A productivity-based test for constraints

3.1 Theory

Credit constraints influence not only whether people start businesses but also how much they will invest in whatever businesses they start. Suppose that family-owned firms produce output using capital k and labor ℓ. Let there be several industries, indexed by j, and several geographic regions, indexed by m. Assume that each household i lives in exactly one geographic region and can operate one or more firms in each industry. I take households' decisions about which industries to enter and how many firms to operate as given exogenously.

Let the production function at date t of a firm operated by household i in industry j be

$$y_{ijt} = f_j(A_{ijt}, k_{ijt}, \ell_{ijt}), \tag{9}$$

where A_{ijt} is a vector that measures household i's ability in industry j at date t and $f_j(\cdot)$ is the common production function of all firms in industry j. I assume the timing of production is: At date $t-1$, the household has some information about what A_{ijt} will be but perhaps is not certain. The household then hires capital k_{ijt} and labor ℓ_{ijt} for use next period. Finally, ability A_{ijt} is realized and production takes place. I assume that households maximize expected profits.

The prices of capital, r, and of labor, w, may vary across regions m. I assume there may be some limit on how much capital a household can borrow or rent, but, for simplicity, I assume there are no limits on hiring of labor. Thus, to maximize expected profits in firm ij, household i solves

$$\max_{k_{ijt},\ell_{ijt}} \mathrm{E}_{t-1}\left[f_j(A_{ijt}, k_{ijt}, \ell_{ijt})\right] - r_{m,t-1}k_{ijt} - w_{m,t-1}\ell_{ijt} \tag{10a}$$

$$\text{subject to} \quad k_{ijt} \leq h(b_{i,t-1}, u_{i,t-1}), \tag{10b}$$

where $b_{i,t-1}$ is the household's wealth at date $t-1$, $u_{i,t-1}$ are unobserved variables affecting the credit limit, $h(b, u)$ is a function specifying the credit limit, and E_{t-1} denotes an expectation taken with respect to the household's information set at date $t-1$.

Letting $\mu_{ijt}r_{m,t-1}$ be the Kuhn-Tucker multiplier on the credit constraint (10b), the first-order necessary condition for an optimal choice of capital is

$$\mathrm{E}_{t-1}\left[\frac{\partial f_j(A_{ijt}, k_{ijt}, \ell_{ijt})}{\partial k}\right] = r_{m,t-1}(1 + \mu_{ijt}). \tag{11}$$

I assume that the production function is separable in labor and capital and that the marginal product of capital is log-linear:

$$\frac{\partial f_j(A_{ijt}, k_{ijt}, \ell_{ijt})}{\partial k} = \hat{f}_j(a_{ijt}, \ell_{ijt}) \cdot \alpha_j \xi_{ijt} k_{ijt}^{\alpha_j \xi_{ij} - 1}, \tag{12}$$

where $A_{ijt} = (a_{ijt}, \xi_{ij})$, $\hat{f}$ is an unknown function, and $\alpha_j \xi_{ij} < 1$.[6] Since the production function is separable, labor will play no further role in the analysis. I include it only to emphasize that the capital share $\alpha_j \xi_{ij}$ is unlikely to be constant across firms because some industries are more capital-intensive and others more labor-intensive.

Substituting the marginal product of capital into the first-order condition, an optimal choice of capital satisfies

$$\mathrm{E}_{t-1}\left[\frac{y_{ijt}}{k_{ijt}}\right] = \frac{r_{m,t-1}}{\alpha_j \xi_{ij}}(1 + \mu_{ijt}). \tag{13}$$

In firms where the credit constraint does not bind, $\mu_{ijt} = 0$ and the expected ratio of output to capital, y_{ij}/k_{ij}, depends only on the interest rate $r_{m,t-1}$ and the production parameter $\alpha_j \xi_{ijt}$. But if μ_{ijt} is positive, it will be a function of variables that affect the credit constraint – in particular, wealth. This result motivates a simple test for credit constraints: Do output-to-capital ratios depend on wealth? [7]

Paulson and Townsend (2004) test for credit constraints in Thailand by, among other methods, tabulating median output-to-capital ratios among entrepreneurs in each quartile of the wealth distribution.[8] If there are no constraints, the median output-to-capital ratios should be equal across the four quartiles. Because such a test does not control for entrepreneurs' industry or place of residence, it implicitly assumes that all entrepreneurs have the same technology and face the

[6] Cobb-Douglas production, $f(A, k, \ell) = Ak^\alpha \ell^\beta$, is a special case.

[7] The output-to-capital ratio must be based on total capital in the business, both owned and rented. It is easy to show that the ratio of output to owned capital must be higher among low-wealth entrepreneurs even under perfect credit markets.

[8] An earlier version of Tejerina (2004) applied a similar test to data from Nicaragua.

same interest rates, or at least that variation in technology and market interest rates is uncorrelated with wealth. Equation (13) shows that if this assumption fails – if interest rates are typically higher in low-wealth regions, or if capital shares $\alpha_j \xi_{ij}$ vary systematically with wealth – output-to-capital ratios can be correlated with wealth even though credit markets may be efficient within each region or industry.

Capital prices can vary substantially within a country. Take the case of Nicaragua. Table 2 shows that in the country's 17 departamentos, which are administrative units somewhat larger than a typical U.S. county, median rental prices for homes range from 12% to 30% of the home's value. Land rents vary from 5% to 20%, and annual compound interest rates on loans from 18% to 80%.

Likewise, Table 3 shows that some Nicaraguan industries employ hardly any capital, while others are much more capital intensive. The median food and beverage manufacturer used just 270 córdobas of capital in 2001, equivalent to slightly more than U.S.$20 at the time, and produced 71 times as much output as it had capital. By contrast, among ground transportation services, such as bus, truck and taxi operators, the median firm had 55,400 córdobas of capital, equivalent to more than U.S.$4,000, and produced only about as much output as it had capital.

Geographic variation in interest rates could be interpreted as a credit market failure, since a free flow of capital would equate interest rates across regions. Barriers to capital flows imply that low-wealth regions will have higher interest rates, all else equal, because their marginal product of capital will be higher. However, geographic barriers to capital flows are a different sort of market failure from the moral hazard and enforcement problems envisioned in most analyses of entrepreneurship. Similarly, while a concentration of wealthy entrepreneurs in capital-intensive industries could signal that credit constraints prevent the poor from raising sufficient capital to pay the fixed costs of entry,[9] other market failures, such as government regulations that limit entry to some industries, could produce a similar relationship between wealth and technology. To determine how much of the variation in output-to-capital ratios comes from inefficient allocation of capital within industries and regions, I construct measures of credit constraints that control for entrepreneurs' industry and location.

[9] See Lloyd-Ellis and Bernhardt (2000) for a theoretical perspective on entry costs and credit constraints, and McKenzie and Woodruff (2003) for empirical evidence.

3.2 Estimating equations

Because $r_{m,t-1}$ and α_j enter equation (13) nonlinearly, it is difficult to control for industry and region with the test in this form. To allow linear controls, begin by taking logs of both sides of (13):

$$\ln\left(y_{ijt}/k_{ijt}\right) = \ln r_{m,t-1} - \ln \alpha_j + \ln\left(1+\mu_{ijt}\right) + \ln e_{ijt} - \ln \xi_{ij}, \quad e_{ijt} \equiv \frac{y_{ijt}/k_{ijt}}{\mathrm{E}_{t-1}\left[y_{ijt}/k_{ijt}\right]}. \tag{14}$$

This equation can be written more clearly by defining a vector x_{ij} of dummy variables indicating the entrepreneur's industry and region, and a vector β_1 containing the values of $\ln r_{m,t-1}$ and $\ln \alpha_j$ for each region and industry, so that

$$\ln\left(y_{ijt}/k_{ijt}\right) = x_{ij}'\beta_1 + \ln\left(1+\mu_{ijt}\right) + \ln e_{ijt} - \ln \xi_{ij}. \tag{15}$$

If credit constraints exist, the Lagrange multiplier μ_{ijt} is an unknown and possibly very complicated function of all the variables that may affect how much the entrepreneur is allowed to borrow, such as wealth $b_{i,t-1}$, and all the variables that affect how much the entrepreneur would like to borrow, such as productivity a_{ijt} and technology $\alpha_j \xi_{ij}$. To estimate this function, it must be simplified in some way. I therefore assume that

$$\mathrm{E}[\ln\left(1+\mu_{ijt}\right) | x_{ij}, b_{i,t-1}] = x_{ij}'\beta_2 + \phi(b_{i,t-1}), \tag{16}$$

where β_2 is a vector of unknown parameters and $\phi(b_{i,t-1})$ is a function of unknown form. (This assumption could be relaxed by allowing x_{ij} to include interactions of industries and regions, given enough observations in each industry-by-region cell.) Then, defining $\beta \equiv \beta_1 + \beta_2$ and $\epsilon_{ijt} \equiv \ln\left(1+\mu_{ijt}\right) - \mathrm{E}[\ln\left(1+\mu_{ijt}\right) | x_{ij}, b_{i,t-1}]$, equation (15) is equivalent to

$$\ln\left(y_{ijt}/k_{ijt}\right) = x_{ij}'\beta + \phi(b_{i,t-1}) + \ln e_{ijt} - \ln \xi_{ij} + \epsilon_{ijt}. \tag{17}$$

Equation (17) says that the mean output-to-capital ratio, given an entrepreneur's region, industry and wealth, consists of an industry- and region-specific intercept $x'_{ij}\beta$ and an unknown function of wealth $\phi(b_{i,t-1})$. Although the initial motivation for the industry- and region-specific intercepts was heterogeneity in technology, the intercept $x'_{ij}\beta$ will also absorb any component of the productivity shock e_{ijt} that is common to all firms in an industry or region, and will control for variation in the mean of $\ln(1+\mu_{ijt})$ across industries and regions.

Consider two entrepreneurs in the same industry, and imagine that the first entrepreneur faces credit constraints while the second does not. Suppose the first entrepreneur has interest rate r_1 and Lagrange multiplier μ_1. If the second entrepreneur faces interest rate $r_2 = r_1(1+\mu_1)$, then according to the first-order condition (13), the second entrepreneur (who has $\mu_2 = 0$) will be induced to choose the same level of capital as the first entrepreneur. Thus the product $r_{m,t-1}(1+\mu_{ijt})$ can be interpreted as a shadow interest rate faced by enterpreneur i: It is the market interest rate that would rationalize i's choice of capital if there were no credit constraints.

The unknown function ϕ measures the shadow interest rates faced by different entrepreneurs. The shadow interest rate faced by a particular entrepreneur i is

$$r_i^{shadow} = r\exp[x'_{ij}\beta_2 + \phi(b_{i,t-1}) + \epsilon_{ijt}], \tag{18}$$

where r is the market interest rate. The error ϵ_{ijt} cannot be observed; it is the difference between the true Lagrange multiplier $\ln(1+\mu_{ijt})$ and its mean conditional on wealth. Furthermore, equation (17) shows that ϵ_{ijt} cannot be distinguished from $\ln e_{ijt}$ or $\ln \xi_{ij}$, so it is impossible to, for example, estimate the variance of ϵ_{ijt}. However, one can find the ratio of the shadow interest rates faced by two entrepreneurs, 1 and 2, who have the same x_{ij} and the same ϵ_{ijt}:

$$\frac{r_1^{shadow}}{r_2^{shadow}} = \exp[\underbrace{(x_{1j}-x_{2j}}_{=0})'\beta_2 + \phi(b_{1,t-1}) - \phi(b_{2,t-1}) + \underbrace{\epsilon_{1jt}-\epsilon_{2jt}}_{=0}] = \exp[\phi(b_{1,t-1}) - \phi(b_{2,t-1})]. \tag{19}$$

Thus, fixing some point b^*, $\exp[\phi(b_i) - \phi(b^*)]$ is the ratio of the shadow interest rate for an entrepreneur i with wealth b_i to the shadow interest rate for an entrepreneur at wealth b^*, holding industry, region and ϵ_{ijt} fixed.

The function ϕ shows how much higher interest rates effectively are, on average, for low-wealth people. A survey cannot measure these shadow interest rates directly by asking entrepreneurs about the interest rates they pay, because credit constraints may not only change the rate an entrepreneur pays but also affect whether she gets any loan at all. Rather, an estimate of ϕ infers differences in shadow interest rates from differences in investment choices.

Banerjee and Duflo (2002) also measure shadow interest rates, using a regression of changes in profits on changes in borrowed capital. This approach requires an instrument for borrowed capital, since firms choose their borrowing endogenously. Banerjee and Duflo find an elegant instrument in a regulatory change that created exogenous variation in access to credit. If all firms have the same technology, this is sufficient to estimate the shadow interest rate. However, if technology is heterogenous, Banerjee and Duflo have a random-coefficients regression in which traditional instruments are invalid.[10]

I attempt to avoid the need for instruments by keeping the endogenous variable k_{ijt} on the left-hand side of equation (17) and projecting the shadow interest rate onto lagged wealth $b_{i,t-1}$, which should be exogenous with respect to the date-t production decision. The disadvantage of my approach is that, although lagged wealth is exogenous with respect to the choice of capital, the variation in lagged wealth is not *ceteris paribus*. Even after controlling for the industry and region dummy variables in x_{ij}, a long list of other factors that affect an entrepreneur's demand for and access to credit – such as productivity a_{ijt} and within-industry technology differences ξ_{ij} – may vary with wealth.

As a consequence, ϕ does not tell how an increase in a given entrepreneur's wealth, holding constant all of her other characteristics, would change her shadow interest rate. All ϕ can reveal is the schedule of shadow interest rates faced by entrepreneurs, *given* the joint distribution of wealth, technology and productivity. Still, ϕ can be used to test for credit constraints, since if there

[10] Banerjee and Duflo (2002) show that the interest rate λ that a firm is willing to pay for additional capital can be found from a regression of the rate of change in a firm's profits, $\Delta \ln \Pi_{it}$, on the rate of change of the firm's borrowed capital, $\Delta \ln k^b_{it}$. They also show that λ is an increasing function of the scale parameter $\alpha_j \xi_{ij}$. If $\alpha_j \xi_{ij}$ is not constant across firms, neither is λ, and the regression should be written $\Delta \ln \Pi_{it} = \beta_0 + (\bar{\lambda} + \lambda_i)\Delta \ln k^b_{it} + u_{it}$. The residual in this regression is $u_{it} + \lambda_i \Delta \ln k^b_{it}$. A valid instrument Z_i must satisfy $\mathrm{E}[Z_i \lambda_i \Delta \ln k^b_{it}] = 0$. This requirement is difficult to meet since a valid instrument must also be correlated with $\Delta \ln k^b_{it}$. Banerjee and Duflo's (2002) instrument is a reform that increased the credit access of firms with a relatively large initial capital stock. Since firms with capital-intensive technology are likely to have a larger initial capital stock, the reform likely increased credit precisely for the firms with high λ_j. This implies the instrumental variable estimator is biased upward.

are perfect credit markets $\mu_{ijt} = 0$ for all entrepreneurs and $\phi(b_{i,t-1})$ must be identically zero. Furthermore, the schedule of interest rates revealed by ϕ is an economically interesting object: It tells how much more productive low-wage entrepreneurs are at the margin, compared with high-wage entrepreneurs. Thus the function ϕ indicates the marginal effect of transferring capital from businesses owned by the rich to businesses owned by the poor, even though it does not indicate the effect of transferring wealth from the rich to the poor.

Procedures for estimating both β and ϕ in partially linear models such as equation (17) are well known. By the assumption in equation (16), $\mathrm{E}[\epsilon_{ijt}|x_{ij}, b_{i,t-1}] = 0$. Assume further that

$$\mathrm{E}[\ln e_{ijt} - \ln \xi_{ij}|x_{ij}, b_{i,t-1}] = \mathrm{E}[\ln e_{ijt} - \ln \xi_{ij}]. \tag{20}$$

This assumption requires, among other things, that heterogeneity in technology within industries, ξ_{ij}, be independent of wealth and that the variance of the productivity shocks e_{ijt} be independent of wealth. I discuss tests of this assumption in Section 4. If this assumption is maintained, equation (17) can be estimated by the method of Robinson (1988).

First, observe that

$$\ln(y_{ijt}/k_{ijt}) - \mathrm{E}[\ln(y_{ijt}/k_{ijt})|b_{i,t-1}] = (x_{ij} - \mathrm{E}[x_{ij}|b_{i,t-1}])' \beta + \ln e_{ijt} - \ln \xi_{ij} + \epsilon_{ijt}. \tag{21}$$

If the conditional expectations $\mathrm{E}[\ln(y_{ijt}/k_{ijt})|b_{i,t-1}]$ and $\mathrm{E}[x_{ij}|b_{i,t-1}]$ were known, β therefore could be estimated by the ordinary least squares regression of $\ln(y_{ijt}/k_{ijt}) - \mathrm{E}[\ln(y_{ijt}/k_{ijt})|b_{i,t-1}]$ on $x_{ij} - \mathrm{E}[x_{ij}|b_{i,t-1}]$. Although the conditional expectations are not known, they can be estimated by a standard nonparametric method such as local linear regression; one local linear regression is required for each variable in the vector x. Robinson (1988) proves that these estimated conditional expectations can be plugged into equation (21) to estimate β by ordinary least squares.[11] Second,

[11] Because nonparametric estimators are imprecise where the density of the regressor is low, some observations must be dropped before estimating β. Robinson's (1988) method requires that observation i be dropped from the regression (21) if the estimated density of wealth at b_i is less than $\bar{f}$, where $\bar{f}$ is a trimming parameter that approaches zero as the sample size increases. The density of wealth can be estimated using a kernel density estimator:

$$\hat{f}(b_i) = \frac{1}{Nh} \sum_{j=1}^{N} K\left(\frac{b_j - b_i}{h}\right),$$

if β were known, the function ϕ could be written as

$$\phi(b_{i,t-1}) = \mathrm{E}[\ln(y_{ijt}/k_{ijt})|b_{i,t-1}] - \mathrm{E}[x_{ij}|b_{i,t-1}]'\beta. \tag{22}$$

The estimate of β from the first step can be plugged into this expression to estimate ϕ:

$$\hat{\phi}(b_{i,t-1}) = \hat{\mathrm{E}}[\ln(y_{ijt}/k_{ijt})|b_{i,t-1}] - \hat{\mathrm{E}}[x_{ij}|b_{i,t-1}]'\hat{\beta}. \tag{23}$$

The intercept in this model is not identified, because adding a constant to $x_{ij}'\beta$ and subtracting the same constant from the function ϕ does not change equation (17). Thus the vector x_{ij} cannot include an intercept, and one category must be omitted from each set of dummy variables. Finally, confidence intervals can be found by bootstrapping, just as for local linear regression.[12]

Using this method, a test for credit constraints can proceed as follows: Estimate equation (17) either with or without the industry- and region-specific intercepts. If the function ϕ depends nontrivially on wealth, entrepreneurs at some wealth levels face credit constraints. Including the intercepts controls for heterogeneity in technology and local interest rates that may be correlated with wealth, as well as common shocks to all firms in an industry or region. Omitting the intercepts is computationally simpler – equation (17) then amounts to a nonparametric regression of $\ln(y_{ijt}/k_{ijt})$ on wealth – but cannot distinguish credit constraints within an industry or region from heterogeneity across industries and regions.

---

where $K(u)$ is a kernel function such as the Epanechnikov kernel. As with the bandwidth, I choose the trimming parameter by visual inspection of the results for different values of $\bar{f}$ and report my choices with the estimates.

[12]Since the intercept is not identified, some normalization must be imposed on the bootstrap estimates of ϕ. Otherwise, the confidence intervals may be too wide or too narrow because the estimated level of ϕ will change with each bootstrap sample according to the mean level of $x_{ij}'\beta$ in that sample. The normalization I impose is

$$\sum_{k=1}^{T} \hat{f}(b_k^*)\hat{\phi}(b_k^*) = \sum_{k=1}^{T} \hat{f}(b_k^*)\hat{\phi}^{(bs)}(b_k^*),$$

where $b_1^*, \ldots, b_T^*$ are the grid points at which ϕ is estimated, $\hat{f}(b_k^*)$ is a kernel density estimate of the density of log wealth at b_k^* computed using the original sample, $\hat{\phi}$ is the estimate of ϕ using the original sample, and $\hat{\phi}^{(bs)}$ is the estimate of ϕ using a bootstrap sample. This normalization holds constant the expectation of $\phi(b)$ relative to the density of wealth in the original sample.

4 Applying the productivity test

As an illustration, I estimate equation (17) on data from Nicaraguan households operating non-farm businesses. I measure wealth in 1998 and the output-to-capital ratio in 2001. Each household reported information about up to two non-farm businesses in 2001. (I do not include farms as businesses because data on the value of farm output in 2001 were incomplete. However, the relationship in equation (17) must hold for data on any subset of industries. This means that, unlike tests of occupation choice, there is no harm in estimating equation (17) on only some kinds of entrepreneurs.) I consider each business as a separate observation but drop all businesses that used zero capital and, as in the probability estimates, all households whose net assets were less than or equal to zero in 1998. Thus the data include 1,291 businesses operated by 1,066 households. Table 4 gives summary statistics.

I begin by estimating equation (17) without the controls for industry and region. Figure 2 shows a local linear regression estimate of the mean of $\ln(y_{ijt}/k_{ijt})$ as a function of wealth for household-owned businesses in Nicaragua. Output-to-capital ratios decrease rapidly with wealth over most of the wealth distribution in Nicaragua.

This relationship could have several causes: capital-intensive industries may have disproportionately many wealthy entrepreneurs; regions with low interest rates may have disproportionately many wealthy households; or, within regions and industries, capital may be allocated inefficiently. To distinguish among these possibilities, I add controls for 17 region categories, one for each departamento, and 106 industry categories, one for each of the four-digit International Standard Industrial Classification categories represented in the data. Figure 3 shows the results of estimating equation (17) with the controls by Robinson's (1988) method.

Including controls for industries and regions dramatically reduces the slope of the relationship between wealth and output-to-capital ratios. Omitting the controls, as in Figure 2 or the tables of Paulson and Townsend (2004), thus may substantially overstate the degree of credit constraints. Nonetheless, Figure 3 shows that output-to-capital ratios do decrease with wealth in Nicaragua even *within* industries and departamentos. If the assumptions of the model are correct, this result indicates that Nicaraguan credit markets do not allocate capital efficiently even among firms in the same industry and region.

Recall that $\exp[\phi(b_2) - \phi(b_1)]$ is the ratio of the shadow price of capital for an entrepreneur with wealth b_2 to the mean shadow price for an entrepreneur with wealth b_1. Figure 4 plots these ratios, taking the median wealth level as the base and using the estimate of $\phi(b_i)$ that includes controls for the entrepreneur's industry and region. Below the median wealth level, it is impossible to reject that the shadow interest rate does not depend on wealth. But the shadow interest rate drops quickly for wealth above the median. For an entrepreneur at the 95th percentile of the wealth distribution, with log wealth in 1998 of 12.45 (equivalent to 255,000 córdobas or about U.S. $25,000), the shadow price of capital is 35% below the shadow price at the median. For an entrepreneur at the 99th percentile of the wealth distribution, with log wealth of 13.49 (equivalent to 722,000 córdobas or about U.S. $70,000), the shadow price of capital is 70% below the shadow price at the median.

4.1 Threats to identification

I assumed above that $\mathrm{E}[\ln e_{ijt} - \ln \xi_{ij} | x_{ij}, b_{i,t-1}] = \mathrm{E}[\ln e_{ijt} - \ln \xi_{ij}]$. This assumption says that 1) the log of the expectations error e_{ijt} is uncorrelated with the regressors; and 2) variation in technology across firms within an industry, ξ_{ij}, is independent of the owner's wealth.

By definition, $\mathrm{E}[e_{ijt} | x_{ij}, b_{i,t-1}] = 0$. This implies $\mathrm{E}[\ln e_{ijt} | x_{ij}, b_{i,t-1}] = \mathrm{E}[\ln e_{ijt}]$ only if the variance of e_{ijt} does not depend on x_{ij} and $b_{i,t-1}$, which is not necessarily true. In principle, this could be tested by nonparametrically estimating the variance of y_{ijt}/k_{ijt} conditional on wealth, industry and region. However, nonparametric estimation is difficult when the set of conditioning variables is high-dimensional, for example if the data include a rich set of industry categories.[13]

It might appear that the second part of the assumption, $\mathrm{E}[\ln \xi_{ij} | x_{ij}, b_{i,t-1}] = \mathrm{E}[\ln \xi_{ij}]$, could be tested given repeated measurements of both wealth and the output-to-capital ratio. Assuming regional interest rates are constant, first-differencing equation (17) would eliminate $\ln \xi_{ij}$ to yield

$$\ln(y_{ij3}/k_{ij3}) - \ln(y_{ij2}/k_{ij2}) = \phi(b_{i2}) - \phi(b_{i1}) + \ln e_{ij3} - \ln e_{ij2} + \epsilon_{ij3} - \epsilon_{ij2}. \tag{24}$$

[13] The homoskedasticity assumption could also be avoided by applying nonlinear least squares to equation (13) using the semiparametric multiplicative model in Robinson (1988), at the cost of much greater computational difficulty.

However, wealth in period 2, b_{i2}, is almost certainly correlated with the productivity shock in period 2, e_{ij2}. Thus an instrument for b_{i2} is required, and equation (24) would have to be estimated by nonparametric instrumental variables as in Newey, Powell and Vella (1999).

An alternative test of the identifying assumptions would estimate $\phi(b_{i,t-1})$ using data on people who are known not to face any credit constraints. The Lagrange multiplier μ_{ijt} should be zero for these people regardless of wealth, so the function ϕ should not depend on wealth. If the estimated ϕ for unconstrained people does depend on wealth, the identifying assumptions are likely invalid; if the estimated ϕ does not depend on wealth, the assumptions may be seen as more plausible.

4.2 Testing the assumptions

The 1998 Nicaragua survey asked each household whether it applied for a loan from any source in the past 12 months, whether it received a loan, and the reason for not seeking a loan if no loan was sought. In addition, each household holding a loan was asked whether it would have liked to borrow more money on the same terms. Table 5 describes these data. Few households appear to have unconstrained access to credit. Only 17 percent of households borrowed money, and half of these would have liked to borrow more money on the same terms but could not do so. About 13 percent of households said they did not need loans. I categorize any household that either held loans and did not want to borrow more money on the same terms, or said it did not need a loan, as unconstrained in 1998. This group constitutes 21% of all households, and 20% of households that owned businesses. I categorize the rest of the households as "possibly constrained."

In results that I do not report here, I find that the relationship between 2001 output-to-capital ratios and 1998 wealth is essentially the same for the unconstrained and possibly constrained groups. This could indicate that the entire test is meaningless, but 1998 financial access could also be a very poor indicator of whether households faced constraints in 2001. Access to and need for credit could change substantially between 1998 and 2001. New investment opportunities could arise, old ones could vanish, bank branches could open and close, and so on. I therefore investigate the relationship between output-to-capital ratios in 1998 and wealth for the two groups.

Pre-1998 wealth data for the Nicaragua survey households are not available, so I now measure both wealth and output-to-capital ratios in the same year. Measuring both variables at the same time could induce a correlation between wealth and productivity even if there are no credit constraints, because families might change their assets in response to productivity shocks. For non-farm businesses in the Nicaragua data, only the most recent two weeks of output are measured, while wealth is measured at the end of the two weeks. The high frequency of the measurement may reduce the bias: Measured wealth – which includes homes, land, business investments and durable goods but not, say, petty cash on hand – is unlikely to change greatly in response to shocks in the past two weeks. For farm businesses, an entire year of output is measured. In case this leads to bias, I compute results separately for non-farm businesses and for all businesses.

Measurement error also can produce spurious correlations between wealth and output-to-capital ratios when these variables are measured simultaneously. Business assets k are probably measured with error. Some business assets are rented; others are owned by the household and form part of household wealth. Suppose that output-to-capital ratios, if measured without error, do not depend on wealth. A positive measurement error in business assets will increase measured wealth and decrease the measured output-to-capital ratio, producing a negative correlation between measured output-to-capital ratios and measured wealth.

Measurement error problems are not readily solved in nonlinear and nonparametric models. However, to determine whether the unknown function ϕ in equation (17) depends on wealth b, it is sufficient to use a linear approximation, $\phi(b) \approx \gamma b$. The coefficient γ can be estimated in a linear regression and, if it is not zero, there are credit constraints. In a linear model, an instrumental variable for wealth eliminates the measurement error problem. A valid instrument is readily available: the value of a family's owned home, household durable goods and bank deposits. This is clearly correlated with total wealth but plausibly uncorrelated with measurement error in business assets.

Let $b_{i,1998}$ be the wealth of family i in 1998. Let x_{ij} be a vector of region and industry dummy variables, as before. Let δ_i be an indicator variable whose value is 0 if family i reports that it was financially unconstrained in 1998, and whose value is 1 if the family is in the constrained category.

Consider the linear regression

$$\log\left(y_{ij,1998}/k_{ij,1998}\right) = c_0 + c_1\delta_i + c_2\delta_i \log b_i, 1998 + c_3(1-\delta_i)\log b_i, 1998 + x'_{ij}\beta + u_{ij}, \tag{25}$$

where u_{ij} is a disturbance with mean zero. As in the partially linear model, $x'_{ij}\beta$ absorbs differences in technology across industries and differences in interest rates across regions, as well as common shocks to all businesses in a region or industry. If there are perfect credit markets for all families, then wealth does not affect optimal investment, so $c_1 = c_2 = c_3 = 0$. However, if families that report being financially constrained indeed face imperfect credit markets, the model predicts that:

- $c_1 > 0$: Firms operating under credit constraints have a higher average marginal product of capital than firms not operating under constraints.
- $c_2 \neq 0$: Constrained families' production decisions depend on wealth.
- $c_3 = 0$: Unconstrained families should show no evidence of constraints.

To test for credit constraints in the regression (25), I first test $c_3 = 0$ against $c_3 \neq 0$. This tests an overidentifying restriction: A rejection of $c_3 = 0$ means that the theory or the identifying assumptions are incorrect, since theory says unconstrained families' decisions should not depend on wealth, and therefore the test must be invalid. If I fail to reject $c_3 = 0$, I will test $c_2 = 0$ against $c_2 \neq 0$. Rejecting $c_2 = 0$ means rejecting the null of perfect credit markets for the households that are possibly constrained.

This regression provides only a statistical test for the existence of credit constraints. If $c_2 = 0$ is rejected, the magnitude of c_2 is not necessarily related to economic parameters of interest such as the magnitude of credit constraints. Indeed, if the null is rejected, there may be selection into constrained or unconstrained status; δ_i may not be independent of ϵ_{ij}; and the coefficients in the regression may not be consistently estimated.

Table 6 reports results for the regression (25). The top panel includes only non-farm businesses that employ physical capital. The bottom panel adds farm businesses to the sample. In both panels, the data are weighted according to the survey sampling weights.

Column 1 of each panel omits industry and geographic fixed effects and does not instrument for wealth. There is a strong negative correlation between wealth and output-to-capital ratios for both possibly constrained and unconstrained families. Column 2 continues to omit the fixed effects but uses non-business assets as an instrument for wealth. This specification also produces a negative correlation between output-to-capital ratios and wealth among the unconstrained when only non-farm businesses are considered. Column 3 of each panel includes fixed effects for departamentos and four-digit industries but does not instrument for wealth. Again, there is a negative correlation between wealth and capital intensity for unconstrained families. Since the correlation should be zero for unconstrained entrepreneurs, omitting either the fixed effects and the instrument apparently produces spurious evidence of credit constraints.

Column 4 includes both the fixed effects and the instrument for wealth. Now the correlation between wealth and output-to-capital ratios is essentially zero for unconstrained families, but it remains negative and statistically significant at least at the 7% level for the possibly constrained families. A variety of robustness checks did not change the qualitative results. Columns 5 and 6 of Table 6 show that the results are robust to using fixed effects for municipios rather than departamentos, or for two-digit rather than four-digit industries. In regressions that I do not report here, I continued to find a large and negative coefficient on wealth for possibly constrained families, and essentially a zero coefficient for unconstrained families, when I omitted sampling weights or measured wealth in levels rather than logs.

The results suggest that when controls are included for industries and regions, and when an instrument is used for wealth if wealth and business assets are measured simultaneously, the output-to-capital ratio regression does not produce spurious evidence of credit constraints among entrepreneurs known to be unconstrained. In particular, the potential bias from measuring wealth and productivity at the same time does not appear to generate incorrect results. This may provide some grounds for confidence in the identifying assumptions. If the identifying assumptions are correct, the results show that in the possibly constrained group – which includes 80% of Nicaraguan entrepreneurs – at least some families indeed face credit constraints.

5 Conclusion

I reviewed two common tests for credit constraints in entrepreneurship: a test examining the probability of starting a business as a function of wealth, and a test examining the relationship between output-to-capital ratios and wealth.

The occupation choice test relies on the assumption that wealth and entrepreneurial ability are unrelated, or that wealth and entrepreneurial ability have a known parametric relationship. These assumptions may be difficult to justify. In any event, the test in practice is extremely sensitive to the definition of a business. In rural Nicaragua, this test showed little evidence of credit constraints when only non-farm enterprises were counted as businesses but strong evidence of credit constraints when some farms also were counted as businesses.

Tests of the relationship between output-to-capital ratios and wealth appear more promising. If the test does not include controls for differences in technology across industries, it can produce spurious or overstated evidence of credit constraints. I showed how to include the necessary controls in a semiparametric model. The procedure, based on the partially linear model of Robinson (1988), measures credit constraints by estimating the shadow interest rates faced by entrepreneurs at different wealth levels.

Applying the output-to-capital test to data from Nicaragua, I found that the implicit interest rate for high-wealth entrepreneurs was at least 35% below the rate at the median wealth level. When I combined the test with data on whether entrepreneurs claimed to face financial constraints, I found no spurious evidence of constraints for the 20% of entrepreneurs identified as unconstrained, but rejected the hypothesis of perfect credit markets for the remaining 80% of entrepreneurs. The results suggest there are substantial credit constraints on entrepreneurship in Nicaragua.

Policymakers aiming to improve credit markets need to know not only whether credit constraints exist but, if so, what market failure causes them. Appropriate policies for one sort of market failure may be useless against other market failures. Neither test I discussed reveals the source of any credit constraints that are found. Paulson and Townsend (2003) test different models of credit constraints using data on occupation choices. Further work is needed to find ways of distinguishing credit market problems without the strong assumptions required in occupation choice models.

Data Appendix

The applied examples in this paper employ cross-sectional and panel data on households, family farms and family-owned businesses from the 1998 and 2001 Living Standards Measurement Surveys conducted by Nicaragua's National Institute for Statistics and Census.

The 1998 survey contains detailed socio-economic and financial data on a stratified random sample of 4,237 households with 23,208 household members from throughout Nicaragua. In addition to basic demographic data, households reported the current value of numerous assets and liabilities and their participation in business, farming and the formal and informal financial sectors.

For 300 households, some or all assets and liabilities were not reported in 1998 – most frequently, the household did not give a value for its home – and I am unable to compute net worth. For an additional 14 households, data on the age of the head of household, financial participation or business ownership are missing. Sampling weights reflecting the stratified survey design are missing for 33 more households. I drop the 347 households with missing data and am left with a sample of 3,890 families in 1998.

A.1 Attrition

The 2001 survey was designed to form a panel with the 1998 survey. However, among the 3,890 households for which there are complete data in 1998, interviewers could find only 2,932, or 75%, in the same census segment in 2001.

To account for the loss of households, I divide the 1998 sample into cells by head's age (three terciles), number of household members (three terciles), wealth (three terciles), education of the most educated adult in the household (above or below median), whether the household operated a farm business in 1998 and whether the household operated a non-farm business in 1998. I then compute the probability within each cell that a household interviewed in 1998 would be found in 2001. I assign to each household found in 2001 an "attrition weight" that is the inverse of this probability. Because the data also include sampling weights reflecting the stratified survey design, I multiply the attrition weights by the sampling weights to produce the correct weights for estimation. I use these combined weights for all estimates that involve 2001 data, and the original sampling weights for the estimates in Tables 5 and 6 that employ only 1998 data. As a robustness check, I constructed alternative weights using different numbers of cells and using a probit model. None of the results were sensitive to using these alternative weights or to omitting them altogether and using only the sampling weights.

Some of the 1998 households split into multiple households by 2001. I treat all of the 2001 households that originated in a single 1998 household as a single family. The 2001 data files also contain observations on families that were added to the survey in that year; because these families cannot be linked to the 1998 data, I drop them from the analysis.

A.2 Wealth

I compute household net worth in 1998 as the sum of the value of owned home, household durable goods, owned business and farm assets, and financial savings, minus household debts to all creditors.

Households reported the current value of any of 25 durable goods that they owned, from clothes irons and toasters to boats and cars. Business assets included inventory, raw materials, vehicles, furniture, machinery, buildings, and land. Farm assets included land, livestock, tractors, tools, silos, chicken coops, and 22 other categories of equipment and buildings. If a household reported that it owned less than 100% of a business, I multiply the value of the business assets by the share that the household owned in computing the contribution to household wealth.

Financial savings included funds deposited in banks and savings cooperatives; with friends, relatives or store owners; or kept in the home. Debts included money owed to banks, finance companies, credit cards, cooperatives, moneylenders, and friends and relatives, as well as trade credit. Households reported whether they had lent money to others but not the amount of these loans, so I could not include loans to others as an asset.

A.3 Capital prices

Homeowners in 1998 reported not only the current value of their homes but also an estimated annual rental price. Similarly, owners of farmland reported what their land would rent and sell for. I use the ratio of rents to sale values of owned homes and land to compute median rental prices for physical capital and land in each of Nicaragua's 17 departamentos. About five percent of households reported an annual rental value for their homes that exceeded the reported sale value of the home. Because this may indicate that either the rental value or sale value was measured with error, I compute the median rental to sale price ratio among households reporting a ratio less than unity. These capital prices are reported in Table 2.

In 2001, homeowners and landowners reported only rental prices, not sale prices, so I cannot estimate capital prices for 2001.

A.4 Businesses

Households in 1998 reported details on up to three non-farm businesses that they owned. Households in 2001 answered questions about up to two businesses. For each non-farm business, households reported the number of months that the business operated out of the past year; the value of output in the most recent two weeks of operation, including output used in the home or bartered; the current value of all owned physical capital; and amounts paid for rented capital in the past year. Capital included land, buildings, vehicles, equipment, unsold inventory and unused raw materials. To compute the capital employed in each business, I divide payments for rented capital by the departamento median rental price of owned homes and then add the value of owned capital. Because I have no data on capital prices in 2001, I continue to use the 1998 capital prices in valuing rented capital in 2001.

I calculate non-farm businesses' output-to-capital ratios on an annual basis, multiply output in the most recent two weeks of operation by the number of weeks per year that the business operates. This accounts for the possibility that seasonal businesses may have high output-to-capital ratios when they are open to make up for the part of the year when the capital is idle.

A.5 Farms

Each household that produced agricultural output in 1998, whether for sale or its own use, reported detailed farming information for the past year. Households reported the quantity produced and amount sold of their 19 most important crops from a list of nearly 300 possibilities; cut wood; five kinds of meat and poultry; and 13 agricultural products, such as milk and eggs. Households also reported the current value of owned land, the amount paid for rented land; and the value of all farm equipment and structures in use, whether owned or rented.

When households sold some of a crop or type of meat, they reported the sale price. I value the unsold output, such as that used in the home, at the same price. I value output of products that the household did not sell according to the national median price in the data for sales of that output. Although it would have been preferable to use prices at the departamento level, there were too few sales of most products to permit this.

Quantities of crops grown were reported in more than 60 different units of measure, some of them not standardized. Where possible, I convert output measurements to pounds before computing the national median prices for valuing unsold output. Where I cannot convert measurements to pounds, I compute a national median price conditional on the units of measure.

For some crops, many households reported output in non-standardized units but no households reported sales measured in these units. For example, hundreds of households grew corn and beans for their own consumption and reported the quantity in "bags." No corn or beans were sold by the bag, and I could not determine any standard in Nicaragua for the size of a bag of corn or beans. In these cases, it is impossible to value the output, and I assign it a value of zero. This primarily affects households engaged in subsistence farming, so it should have limited impact on the analysis of the business decisions of larger farms.

To compute the total capital used in a farm, I divide payments for rented land by the departamento median rental price of farmland, then add the value of owned land and the value of owned and rented structures, equipment and livestock.

References

Aghion, Philippe, and Patrick Bolton, 1997, "A Theory of Trickle-Down Growth and Development," *Review of Economic Studies*, 64(2), 564–591.

Banerjee, Abhijit V., and Esther Duflo, 2002, "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program," manuscript, Massachusetts Institute of Technology.

Banerjee, Abhijit V., and Andrew F. Newman, 1993, "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101(2), 274–298.

Blanchflower, David G., and Andrew J. Oswald, 1998, "What Makes an Entrepreneur?" *Journal of Labor Economics*, 16(1), 26–60.

Buera, Francisco J., 2003a, "A Dynamic Model of Entrepreneurship with Borrowing Constraints," manuscript, University of Chicago.

Buera, Francisco J., 2003b, "Identification and Testable Implications of Occupational Choice Models," manuscript, University of Chicago.

Evans, David S., and Boyan Jovanovic, 1989, "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy*, 97(4), 808–827.

Fan, Jianqing, 1992, "Design-adaptive Nonparametric Regression," *Journal of the American Statistical Association*, 87(420), 998–1004.

Fazzari, Steven M., R. Glenn Hubbard, and Bruce C. Petersen, 1988, "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*, 1988(1), 141–195.

Galor, Oded, and Joseph Zeira, 1993, "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60(1), 35–52.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S. Rosen, 1994a, "Entrepreneurial Decisions and Liquidity Constraints," *RAND Journal of Economics*, 25(2), 334–347.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S. Rosen, 1994b, "Sticking it Out: Entrepreneurial Survival and Liquidity Constraints," *Journal of Political Economy*, 102(1), 53–75.

Hurst, Erik, and Annamaria Lusardi, 2004, "'Liquidity Constraints, Household Wealth, and Entrepreneurship," *Journal of Political Economy*, 112(2), 319–347.

Jeong, Hyeok, and Robert M. Townsend, 2003, "Growth and Inequality: Model Evaluation Based on an Estimation-Calibration Strategy," manuscript, University of Southern California and University of Chicago.

Kaplan, Steven N., and Luigi Zingales, 1997, "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?" *Quarterly Journal of Economics*, 112(1), 169–215.

Lloyd-Ellis, Huw, and Dan Bernhardt, 2000, "Enterprise, Inequality and Economic Development," *Review of Economic Studies*, 67(1), 147–168.

Matzkin, Rosa L., 1992, "Nonparametric and Distribution-Free Estimation of the Binary Threshold Crossing and the Binary Choice Models," *Econometrica*, 60(2), 239–270.

McKenzie, David J., and Christopher Woodruff, 2003, "Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises," manuscript, Stanford University and University of California, San Diego.

Newey, Whitney K., James L. Powell, and Francis Vella, 1999, "Nonparametric Estimation of Triangular Simultaneous Equations Models," *Econometrica*, 67(3), 565–603.

Paulson, Anna L., and Robert M. Townsend, 2003, "Distinguishing Limited Commitment from Moral Hazard in Models of Growth with Inequality," manuscript, Federal Reserve Bank of Chicago and University of Chicago.

Paulson, Anna L., and Robert M. Townsend, 2004, "Entrepreneurship and financial constraints in Thailand," *Journal of Corporate Finance*, 10(2), 229–262.

Robinson, P.M., 1988, "Root-N-Consistent Semiparametric Regression," *Econometrica*, 56(4), 931–954.

Roy, A.D., 1951, "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers, new series*, 3(2), 135–146.

Tejerina, Luis, 2004, "Entrepreneurship and Financial Constraints: The Case of Nicaragua," manuscript, Inter-American Development Bank.

Table 1: Household business ownership and wealth.

	Urban	Rural
Full sample		
Households with complete data in 1998	2120	1770
Households still present in 2001	1620	1312
Weighted fraction operating non-farm business in 1998	0.45	0.22
Weighted fraction operating any business in 1998	0.49	0.51
1998 net assets, córdobas:		
median	20,000	10,680
mean	61,868	32,629
standard deviation	240,159	78,247
Households without non-farm business in 1998		
Households still present in 2001	852	1008
Weighted fraction operating non-farm business in 2001	0.35	0.20
1998 net assets, córdobas:		
median, households with business in 2001	16,030	12,070
median, households without business in 2001	14,450	9,455
median, all households	15,150	10,470
mean, all households	53,820	32,333
standard deviation, all households	227,647	80,875
Households without any business in 1998		
Households still present in 2001	775	542
Weighted fraction operating any business in 2001	0.37	0.53
1998 net assets, córdobas:		
median, households with business in 2001	15,600	8,160
median, households without business in 2001	13,730	3,765
median, all households	13,786	5,500
mean, all households	42,953	16,261
standard deviation, all households	124,215	44,023

Data used for estimating probability of starting a business as a function of wealth. Net assets include value of home, household durable goods, business and farm assets including livestock, and financial savings, minus debts. In 1998, one U.S. dollar was worth approximately 10 córdobas. All statistics except sample sizes computed with sampling weights adjusted for attrition.

Table 2: Median capital prices by departamento, 1998.

Region Departamento	Rental price (%) Home	 Land	Annual interest (%) Simple	 Compound
Managua				
Managua	16	7	60	80
Pacífico				
Carazo	12	10	36	43
Chinandega	23	8	48	60
Granada	12	5	36	43
León	14	7	42	51
Masaya	12	5	36	43
Rívas	15	8	36	43
Central				
Boaco	17	8	36	43
Chontales	16	6	24	27
Estelí	15	10	36	43
Jinotega	12	13	32	35
Madriz	30	20	18	18
Matagalpa	12	10	24	27
Nueva Segóvia	12	10	36	43
Atlántico				
raan	24	17	30	30
raas	18	13	28	30
Río San Juan	24	17	24	27
Nicaragua				
National median	15	10	36	43
Observations	2836	821	521	521
Test for equality of medians:				
$\chi^2(16)$	188.87	68.04	29.15	29.15
p-value	0.0000	0.0000	0.0229	0.0229

Rental price of home is ratio of annual rental value to sale value reported by homeowners, dropping 157 households that reported rental value larger than sale value. Rental price of land is ratio of annual rental value to sale value of land reported by landowners, dropping 96 properties where landowner reported rental value larger than sale value. Annual simple interest and annual compound interest are computed from rates reported by borrowers for loans from all sources with terms of at least one month.

Table 3: Family non-farm businesses' capital and output by two-digit ISIC industry, 2001.

Industry	Businesses: Total	Businesses: No capital	Total Capital[1,2]: Mean	Total Capital[1,2]: Median	Mean rented capital (%)[2]	Mean output[1]	Median Y/K[2]
Natural resources							
Agricultural	8	0	19948	2600	54	15046	1.93
Other mining	5	1	33967	51600	0	94233	3.29
Forestry	3	0	4373	236	0	4609	3.99
Fishing	19	1	28727	2000	42	21331	5.85
Non-ferrous minerals	7	0	1400	500	0	20604	26.00
Manufacturing							
Publishing	2	0	124228	149420	15	316359	1.12
Machines	2	0	11300	15250	24	16167	1.14
Clothing	75	3	3361	1700	0	12723	3.43
Chemicals	1	0	4320	4320	0	15600	3.61
Metal products	13	0	15824	3250	29	33137	4.19
Miscellaneous	36	1	12764	1800	1	60279	8.67
Wood products	16	1	4098	1210	14	32385	10.40
Leather products	9	0	6120	3760	27	59588	10.90
Textiles	17	0	9404	1700	6	53300	24.58
Other mineral products	11	1	2906	225	48	12138	30.59
Basic metals	1	0	220	220	0	13000	59.09
Food and beverages	89	8	2598	270	34	25660	71.31
Sales							
Vehicles and fuel	32	1	40370	7150	6	59063	4.26
Wholesale	34	7	26425	6500	30	185704	6.15
Retail	643	57	11766	2200	19	53020	11.64
Services							
Ground transportation	55	1	86708	55400	40	75204	0.98
Recreation, culture	13	1	38915	9150	19	51859	2.19
Machinery rental	4	0	14826	12600	0	26309	3.47
Social and health	19	5	35168	200	25	39374	5.42
Travel, cargo agencies	6	3	1705	1500	0	62558	6.93
Telecommunications	1	0	730	730	0	5200	7.12
Sanitation	1	0	200	200	0	1517	7.58
Education	12	3	12566	457	83	24665	9.75
Professional	27	2	25274	4600	20	59889	12.13
Construction	99	11	1667	1000	0	25803	13.00
Laundry, barbers, funerals	173	123	1844	220	12	7844	25.28
Hotels and restaurants	102	7	10557	450	11	54310	42.76
Information processing	3	1	191	250	0	4764	52.00
Associations	3	0	9937	850	2	56577	62.86
Real estate	2	2	-	-	-	21585	-
Total	1543	240	15240	1700	25	45056	9.77

Means and medians computed with sampling weights adjusted for attrition. [1]In córdobas. [2]Among businesses with positive capital.

TABLE 4: Entrepreneurs' output-to-capital ratios and wealth.

Businesses	1291
Households	1066
Departamentos	17
Four-digit ISIC industries	106
1998 household net assets, córdobas	
median	21,320
mean	52,995
standard deviation	108,420
2001 business output-to-capital ratio	
median	9.77
mean	67.56
standard deviation	267.17

Data used for estimating mean of 2001 log output-to-capital ratio as a function of 1998 wealth. Businesses with zero capital and households with 1998 assets less than or equal to zero were dropped. Means, medians and standard deviations computed with sampling weights adjusted for attrition.

Table 5: Household financial participation, 1998.

	Percent of households
Borrowing[1]	
Did not apply for a loan because ...	
Loans not offered in community	4.7
Did not know lenders	0.8
Had too much debt already	0.2
Too many costly requirements	2.6
Did not know how to apply for a loan	2.8
Lacked collateral	18.1
Feared losing collateral	6.4
Lacked sufficient or stable income	10.9
Interest rate very high	7.1
Preferred to work with own resources	7.8
Did not have investment opportunity	4.3
Did not need loan	12.6
Other	1.7
Sought loan but received none	2.6
Held loan and ...	
wanted to borrow more on same terms	8.5
did not want more on same terms	8.6
Number of loans held[1]	
None	82.8
1	14.3
2	1.4
3	1.1
4 or more	0.4
Sources of credit[2]	
Bank, finance company, credit card	5.8
Cooperative	3.2
Non-governmental organization	1.9
Moneylender	1.3
Friend, relative, neighbor	5.4
Purchases on credit	9.0

Percentages computed using sampling weights. [1]Not including purchases on credit. [2]Percent of households reporting use of institution in the past 12 months. Households can report use of multiple institutions.

Table 6: Correlation between output-to-capital ratios and wealth, 1998.

Log of output-to-capital ratio, non-farm businesses						
Variable	OLS (1)	IV (2)	OLS (3)	IV (4)	IV (5)	IV (6)
log wealth	-0.270	-0.184	-0.171	-0.105	-0.115	-0.159
-constrained	(0.055)	(0.055)	(0.055)	(0.058)	(0.064)	(0.050)
p-value	*0.0000*	*0.0009*	*0.0023*	*0.0713*	*0.0743*	*0.0016*
log wealth	-0.286	-0.171	-0.116	0.029	0.024	0.060
-unconstrained	(0.077)	(0.109)	(0.084)	(0.091)	(0.094)	(0.087)
p-value	*0.0003*	*0.1176*	*0.1698*	*0.7514*	*0.8014*	*0.4940*
constrained	0.158	0.521	0.718	1.574	1.637	2.511
	(1.003)	(1.316)	(0.985)	(1.065)	(1.103)	(1.035)
constant	4.834	3.611	0.210	0.341	-1.002	0.606
	(0.830)	(1.174)	(1.021)	(1.069)	(1.026)	(1.022)
fixed effects						
geographic	-	-	departamento	departamento	municipio	departamento
industry	-	-	4-digit	4-digit	4-digit	2-digit
businesses	1147	1147	1147	1147	1147	1147
households	959	959	959	959	959	959
R^2	0.06	0.06	0.31	0.31	0.38	0.21

Log of output-to-capital ratio, all businesses						
Variable	OLS (1)	IV (2)	OLS (3)	IV (4)	IV (5)	IV (6)
log wealth	-0.230	0.270	-0.207	-0.071	-0.082	-0.104
-constrained	(0.044)	(0.059)	(0.033)	(0.035)	(0.040)	(0.033)
p-value	*0.0000*	*0.0000*	*0.0000*	*0.0429*	*0.0406*	*0.0020*
log wealth	-0.417	-0.056	-0.247	-0.004	-0.036	0.023
-unconstrained	(0.088)	(0.143)	(0.073)	(0.084)	(0.081)	(0.081)
p-value	*0.0000*	*0.6989*	*0.0008*	*0.9584*	*0.6600*	*0.7783*
constrained	-2.317	-3.432	-0.344	0.868	0.648	1.533
	(1.047)	(1.683)	(0.814)	(0.961)	(0.939)	(0.945)
constant	5.616	1.754	3.590	0.660	-0.414	0.213
	(0.939)	(1.571)	(1.165)	(1.156)	(1.059)	(1.058)
fixed effects						
geographic	-	-	departamento	departamento	municipio	departamento
industry	-	-	4-digit	4-digit	4-digit	2-digit
businesses	2029	2029	2029	2029	2029	2029
households	1723	1723	1723	1723	1723	1723
R^2	0.04	-	0.59	0.58	0.62	0.53

Dependent variable is log ratio of annual business output to capital employed in business. For farms, capital includes value of owned and rented land. All variables measured in 1998. Heteroscedasticity-robust standard errors, computed with clustering, are in parentheses. P-values are for test of the null hypothesis that the coefficient is zero. All regressions in this table were computed with sampling weights. The instrument for wealth in columns 2, 4, 5 and 6 is log(1+value of home, household durable goods and financial deposits).

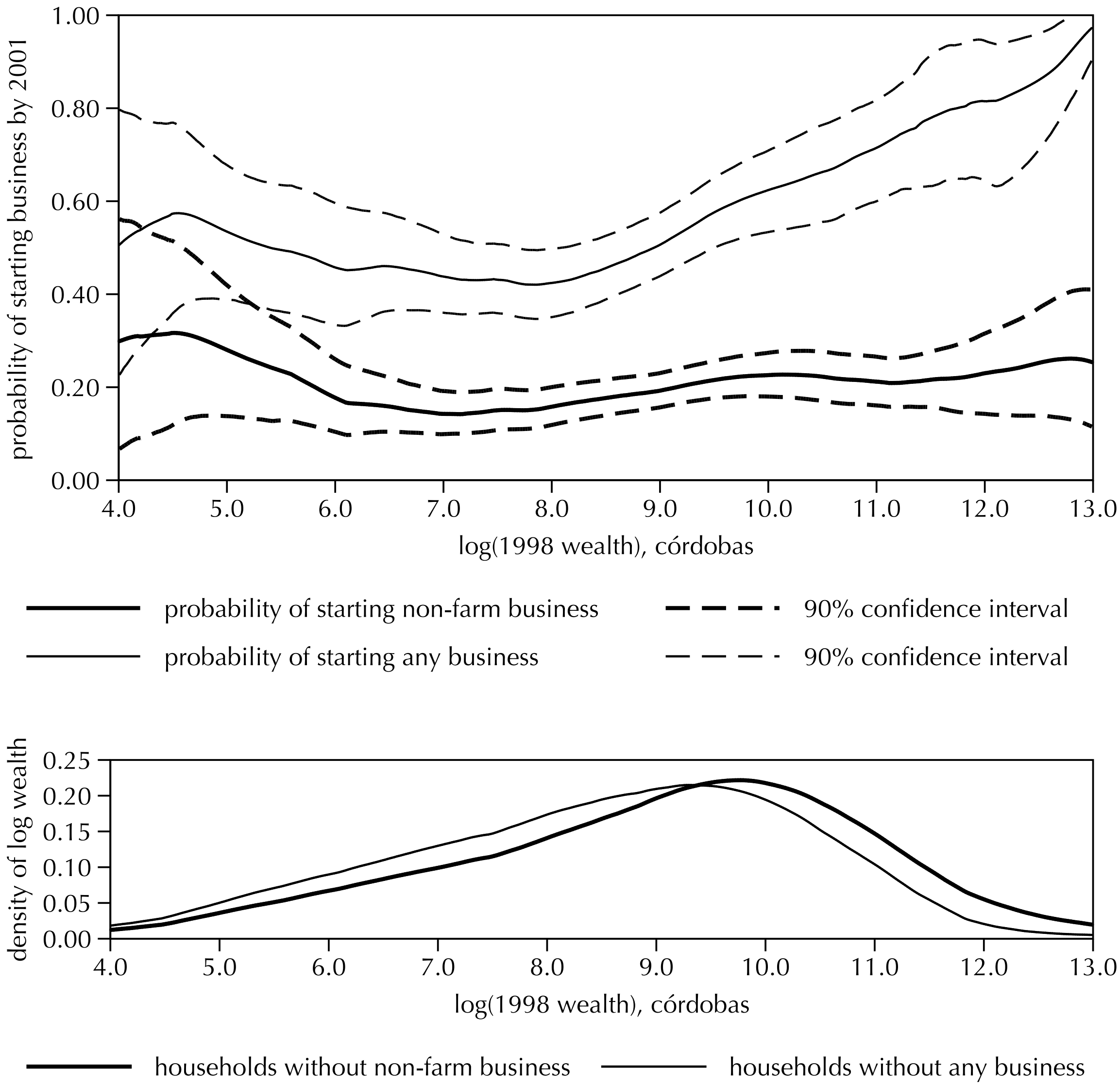

Figure 1: Top: probability of starting a business in rural Nicaragua as a function of household wealth. Thick lines show local linear regression estimate of the probability of owning a non-farm business in 2001, conditional on not owning a non-farm business in 1998 (sample size 1,008); thin lines show estimated probability of owning any business in 2001, conditional on not owning any business in 1998 (sample size 542). Broken lines indicate 90% confidence intervals, constructed by bootstrapping using 10,000 random samples of the same size as the original sample, drawn from the original sample with replacement. Bottom: kernel density of wealth distribution conditional on not owning a non-farm business or any business in 1998. All estimates use Epanechnikov kernel and bandwidth of 1.5 in units of log wealth.

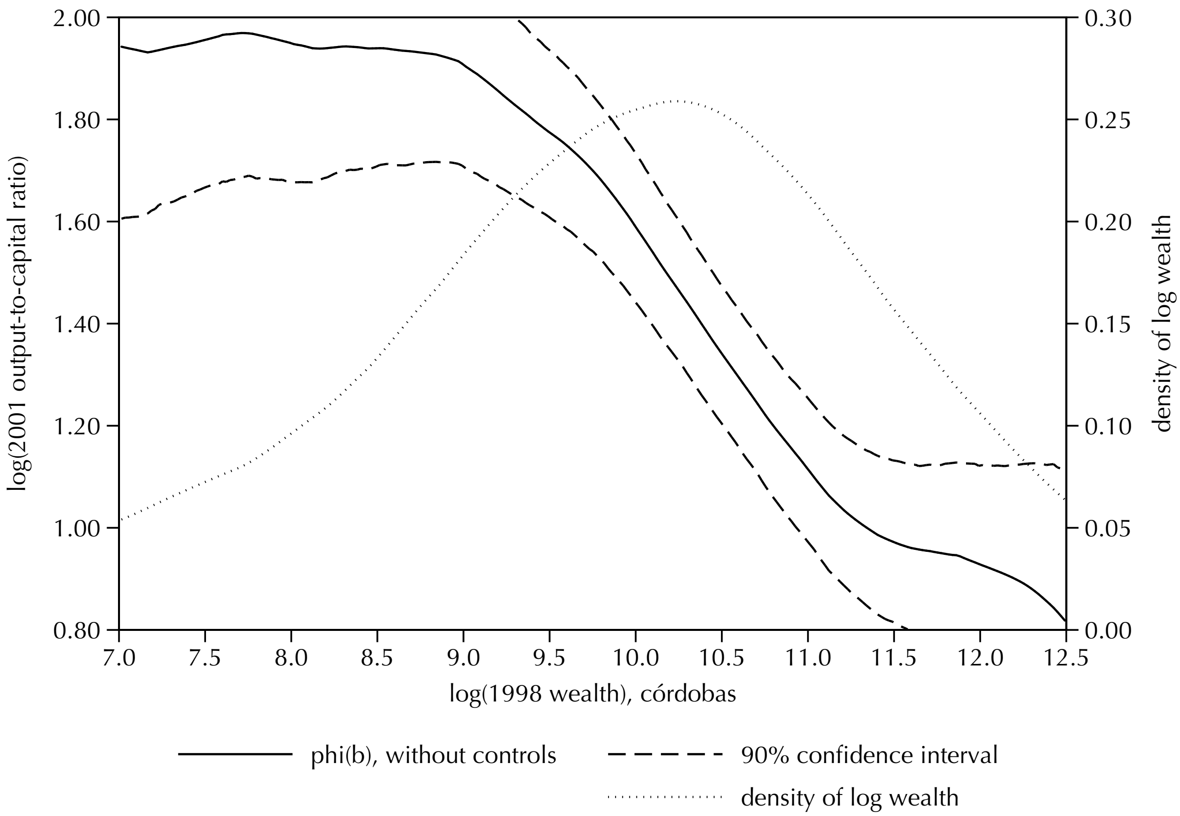

Figure 2: Local linear regression estimate of relationship between entrepreneurs' output-to-capital ratios and wealth in Nicaragua, without controlling for heterogeneity. Sample size is 1,291. Horizontal axis runs from 6.5th percentile to 96th percentile of wealth distribution. Estimate uses Epanechnikov kernel and bandwidth of 1.5 in units of log wealth. Confidence interval constructed by bootstrapping using 3,000 random samples of the same size as the original sample, drawn from the original sample with replacement.

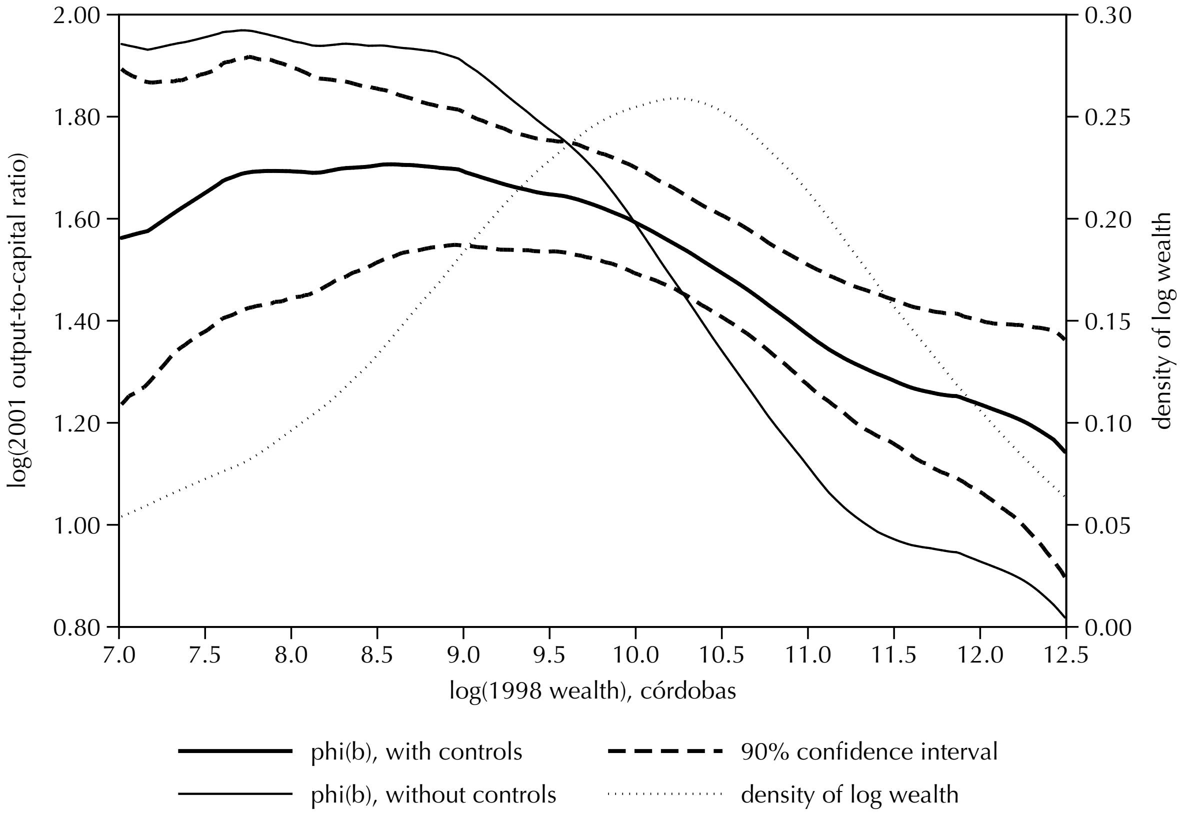

FIGURE 3: Relationship between entrepreneurs' output-to-capital ratios and wealth in Nicaragua, after controlling for heterogeneity. Thick line shows estimated conditional mean of $\ln(y/k)$ given wealth, minus industry- and region-specific intercepts, using Robinson's (1988) method. Sample size is 1,291. Horizontal axis runs from 6.5th percentile to 96th percentile of wealth distribution. Estimate uses Epanechnikov kernel, bandwidth of 1.5 in units of log wealth, and trimming parameter of 0.01. Confidence interval constructed by bootstrapping using 3,000 random samples of the same size as the original sample, drawn from the original sample with replacement. Thin line shows conditional mean without industry and region controls from Figure 2.

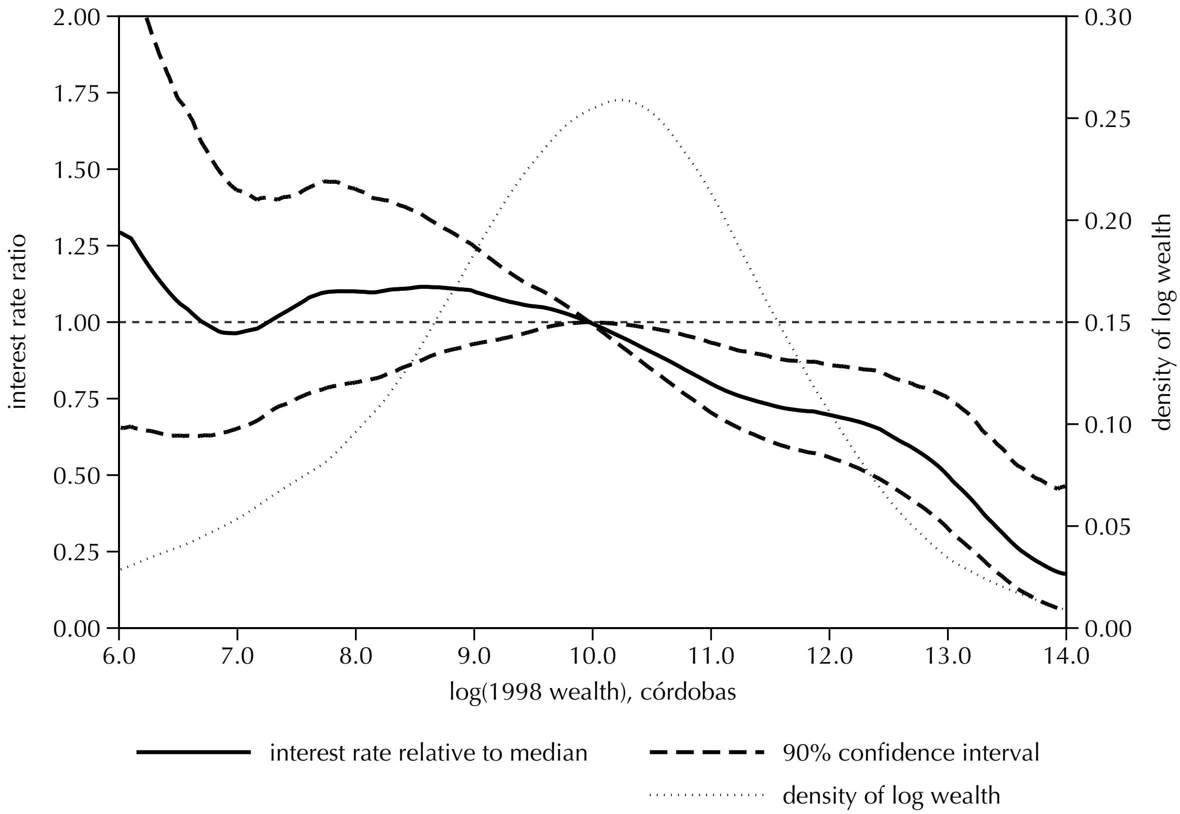

FIGURE 4: Shadow interest rates in Nicaragua as a function of wealth, relative to shadow interest rate at median wealth. Sample size is 1,291. Horizontal axis runs from 2.5th percentile to 99.9th percentile of wealth distribution. Confidence interval constructed by bootstrapping using 3,000 random samples of the same size as the original sample, drawn from the original sample with replacement.