ENTREPRENEURSHIP, ENTERPRISE START-UP, AND THE INVESTMENT CLIMATE IN RURAL AREAS

Wim P.M. Vijverberg
Ph.D. Program in Economics
City University of New York – Graduate Center
365 5th Avenue
New York, NY 10016-4309

December 2008

Abstract: This paper uses data from rural surveys in Nicaragua and Sri Lanka that are specially designed to measure the many facets of the investment climate and their effect on enterprise performance and entrepreneurship selection. In regard to entrepreneurship selection, in Nicaragua, household characteristics are more important than investment climate features; in Sri Lanka, both factors are roughly equally important. The model indicates no role for a latent community factor in Nicaragua, but in Sri Lanka the latent factor absorbs one third of the explanatory power of the community determinants. The paper also examines enterprise start-up. As entrepreneurship is only one of several possible options for a rural household, the same model is also estimated for farming, wage employment, and involvement in nonfarm activities broadly defined.

Keywords: Nonfarm entrepreneurship, Enterprise start-up, Household activity choice, Investment climate, Rural development

JEL Code: O1, J22, R11

Acknowledgement: Research for this paper was done while the author was a consultant for the Agriculture and Rural Development Department of the World Bank. This paper is an updated version of World Bank (2008). Research assistance of Yyannu Cruz Aguayo and Isabel Rodriguez Tejedo and the input of Kees van der Meer, Richard Burcroft and John Lamb are gratefully acknowledged. Opinions expressed in this paper are the author's alone.

1. THE IMPORTANCE OF RURAL ENTREPRENEURSHIP

Across the developing world, rural areas exhibit a richly varied pattern of activities. While, of course, many households derive their livelihood from farming, people are also active as traders, service providers and manufacturers. In fact, rural nonfarm enterprises provide 30 to 45 percent of rural incomes across the developing world and an even higher share for the rural poor (Haggblade et al. 2007).

The level of activity in rural nonfarm enterprises (RNFEs) not only depends on growth in agricultural production and rural consumption but in turn also stimulates it, as RNFE efficiency lowers transaction costs of bringing farm product to consumers and distant markets and of bringing inputs and non-tradable consumer goods and services from distant markets to rural residents (Haggblade et al. 1989; Lanjouw and Lanjouw 2001). Yet, it would be a mistake to think of RNFEs only in terms of a loop between agricultural production and rural consumption. The low cost of rural labor may induce some urban entrepreneurs to establish rural branches, and also lead some rural entrepreneurs to produce for urban markets (Hayami 1998; Haggblade et al. 2007).

In entrepreneurial decisions, it is not only the cost of rural labor and the access to start-up capital that matter. Many facets of the rural investment climate may play a role: the cost of finance, the quality of the rural labor force, the reliability of electricity provision, access to clean water, the communication network (telephone, internet), the cost of transportation, accessibility of the community over all four seasons, the integrity of local government officials, the burden of taxation, the red tape of registration and licensing, the level of business expertise in the community. Indeed, there are many ways in which policy makers can influence rural economic activity—and there are good reasons why they should: employment growth, poverty reduction, spatial population patterns.

This paper delves into the question why households engage in the operation of a RNFE and which households choose to do so. Is it only a matter of skills and resources available within the household? Does a favorable investment climate attract households to entrepreneurship? Data from Rural Investment Climate (RIC) surveys in Sri Lanka and Nicaragua are used to shed light on this question. These new surveys are dedicated to the documentation of the investment climate in rural areas.

When entrepreneurship is defined as the current operation of an enterprise, it is a "stock" concept (Van der Sluis, Van Praag and Vijverberg 2005). But entrepreneurship may also be thought of in dynamic terms: finding market niches, organizing efficient methods of production, exploring new opportunities. If entrepreneurship means the pursuit of success, some ventures must inevitably fail. Indeed, turnover among small enterprises in most countries around the world is often high, with 15 to 20 percent of business activities being abandoned each year—but about an equivalent number is starting up (e.g., Vijverberg 2006).[1] The RIC surveys do not permit an analysis of enterprise deaths, but an analysis of enterprise start-up is feasible. The obvious policy questions about this "flow" concept of entrepreneurship are: does the investment climate have any impact on a household's decision to start up a new enterprise?

[1] Haggblade, Hazell and Reardon (2007) mention a range of 20-50 percent.

Entrepreneurship is just one of several forms of employment and income generation. Households may also pursue wage employment or farm activities. It is useful to analyze these three together, as a transition from a traditional agrarian society to a modern economy with activity in manufacturing, trade and services produced in large-scale enterprises goes through a phase where non-farm self-employment and small-scale entrepreneurship are important. As enterprises become successful, they expand and hire workers who, in an earlier phase of the transition, might have set up their own businesses. Features of the investment climate may encourage entrepreneurship in one society, drawing people out of farming, and may discourage it in another, drawing people into wage employment at larger enterprises.

The next section highlights patterns of employment and income among rural households surveyed in Nicaragua and Sri Lanka. Section 3 addresses econometric issues related to random effects and sample stratification and clustering. Section 4 examines household involvement in non-farm enterprises through regression analysis. Section 5 broadens the perspective to include other forms (wage employment, farming) of income-earning activities. Section 6 investigates enterprise start-up. Section 7 summarizes and concludes.

2. ECONOMIC ACTIVITIES IN RURAL SRI LANKA AND NICARAGUA

The RIC Surveys

In 2003, the World Bank administered its first Rural Investment Climate (RIC) survey. It was an extension of similar survey efforts among urban enterprises over the previous several years (e.g., Batra, Kaufmann and Stone 2003; Dollar, Hallmark-Driemeyer and Mengistae 2005; Bigsten and Soderbom 2006),[2] but these surveys were not shedding light on the conditions in rural areas where after all three quarters of the world's poor reside. Six pilot surveys were launched, two of which are providing data for this study: Sri Lanka in 2003, and Nicaragua in 2004.

The "investment climate" consists of the totality of factors that provide for a reasonable rate of return to investment as perceived by potential private investors (World Bank, 2006:11). Thus, RIC surveys must collect a broad range of information on the environment within which enterprises make economic decisions. As such, not only should entrepreneurs be able to voice their opinion about the investment climate; it is crucial that the survey also gathers objective information about conditions in the community.

Measurement of the rural investment climate and its impact is challenging: (*i*) the investment climate is a multifaceted concept; (*ii*) in rural space there are few large enterprises and myriads of small ones; (*iii*) few enterprises are registered so that a survey cannot rely for its sampling frame on rural lists of enterprises; and (*iv*) enterprise turnover is high, in part because of seasonality driven by agriculture. Therefore, RIC surveys must focus on both enterprise performance and entrepreneurship selection. In line with these objectives, RIC surveys gather data on sampled enterprises (of any size), on households that own them,[3] on a complementary

[2] See http://iresearch.worldbank.org/InvestmentClimate/ for these data.

[3] In these pilot surveys, households owning standalone enterprises were not surveyed. Moreover, enterprises do not have to be owned by a household, in which case the RIC survey would not be able to interview the enterprise owner. However, such cases are actually rather rare.

sample of households that do not own an enterprise, and on the communities where these enterprises and households reside. The focus of the questionnaire is non-farm activities, and only summary information is collected about farming and wage employment. The Sri Lanka survey counts 1123 households and 1327 enterprises in 151 communities. The Nicaragua RIC consists of 1535 households, 1482 enterprises and 97 communities.[4] Sampling weights are applied in the analysis of these data.

Employment

As mentioned in Section 1, entrepreneurship (or self-employment) is just one of the options to earn a living. To get a sense of the importance of entrepreneurship in rural life, Table 1 shows involvement in major economic activities such as self-employment, wage employment and farming. The first three columns summarize activity by household, as judged by employment of household members (column 1) or income received (columns 2 and 3). Columns 4 to 6 reflect economic activity among the adult rural population.

In Nicaragua, 22.7% of rural households are associated with a non-farm enterprise, which engages 14.4% of the adult rural population. The disparity between these percentages is indicative of the diversification of household activity: 35.2% of the households have members engaged in different activities, but only 14.6% of adults actually do so. Only one of every three rural households with an enterprise specializes in it; in particular, many combine it with wage employment. Judged by the income actually received, 22.2% receive at least some amount from their enterprise, and for 22.0% it is greater than a trivial amount of $60 per year.

By comparison, 41.7% of the adult population worked for a wage[5] during the 12 months prior to the survey, residing in 66.4% of the households, and 25.1% worked on their farm, residing in 35.1% of the households. Over the course of a year, many work both on their own farm and for someone else for a wage, especially among males.

In Sri Lanka, non-farm enterprises are more widespread: 29.4% of households are affiliated with at least one enterprise, employing 13.5% of the adult rural population. Among household members, 37% worked for a wage, residing in 76.7% of households, and 17.9% worked on their farm, residing in 41.6% of households. Once again, at the household level, diversification is in full view (43.3%); enterprises are more frequently combined with wage employment than with farming. However, household members usually specialize: only 10.5% of adults engaged in more than one activity during the 12 months prior to the survey—and if they do anything else, it is usually farm work.

According to Table 1, supposedly no Sri Lankan households derive income solely from their non-farm enterprise. This is actually an artifact of the questionnaire, which asked for non-farm sources other than wage earnings only if the household had at least one wage-employed member. This is of minor consequence to the purpose of this study, but the percentages in columns 2 and 3 of Table 1 therefore understate the importance of non-farm enterprises in rural Sri Lanka. Thus, *at least* 18% of the households receive some income from their enterprise; the actual percentage must be higher in light of the fact that nearly 30% of households are affiliated with a non-farm enterprise (column 1).

[4] As pilot surveys of the RIC assessment project, they are not exactly identical in the questions they asked or the population they covered, but for the purpose of this paper they provide data of a breadth in scope that is unmatched. For more information about RIC surveys, see World Bank (2006, 2007).

[5] It is not known whether wage employment was agricultural or non-agricultural.

In both countries, activity diversification among entrepreneurial households is greater than among the rural population overall: roughly two thirds of the households with an enterprise are also involved in other activities. Individuals who work in a household enterprise are also more likely to engage in several modes of employment over the course of a twelve-month period. This parallels findings in many countries around the world (e.g., Moock et al. 1990; Lanjouw 2001; Reardon et al. 2001; Vijverberg 1990, 1998, 1999, 2005).

Both countries reveal interesting gender patterns. In Nicaragua, about 30% of the employed women work in the household's enterprise (15.0% of the adult population, relative to 49.6% of women who work in any capacity), which is substantially higher than the 16.2% of men who do so (14.2% relative to 87.0%). In Sri Lanka, the discrepancy is not quite as large: 25.8% for women (10.0% relative to 38.8%) as compared with 22.2% for men (16.8% relative to 76.3%), but more men (69.1%) than women (54.6%) find wage employment if they work. Thus, in both countries, nonfarm enterprises are important for the employment of women.[6]

Income

Column 2 of Table 1 shows incidence of enterprise income receipt but not quantity. In Nicaragua, 22 percent of total household income comes directly from nonfarm enterprises.[7] Wage income accounts for 49%, net farm income for 11%, remittances for 11%, and other sources contribute 7%. The survey did not distinguish whether wage workers were employed in agriculture or by a non-farm enterprise, but at least a portion of wage income also derives from work in nonfarm enterprises. Figure 1 decomposes household income for each quintile. Among the poorest households, enterprises contribute only 6.8% of their total income. This grows to about 13% in the three middle quintiles, and rises to 28.5% among households in the richest quintile. This positive association between the enterprise income share and the total income amount is quite typical in developing countries (e.g., Reardon 1997).[8] It is open to debate whether this association might reflect access to markets and infrastructure rather than a natural differentiation among households, and whether wage income will overtake both farm and nonfarm enterprise income as the rural economy grows further (Reardon et al. 2001). Descriptive analysis cannot answer such questions; econometric modeling is needed.

3. ECONOMETRIC MODELING

In the econometric analysis below, entrepreneurship is viewed as an outcome at the household level. This avoids the deep econometric complications of a collective household model of activity choices of its members that must allow for two crucial features: interactive decision making among a number of members that varies between households, and a mixture of individual labor market opportunities and joint labor market venture that members might pursue

[6] For both men and women, the percentages of people involved in nonfarm self-employment are somewhat higher than those reported generally for Latin America and the Caribbean area and for South Asia in World Bank (2007, Ch.9). It is possible that the focus of the RIC survey contributes to a more precise reporting of entrepreneurial activities, which are sometimes secondary activities in a household and therefore undercounted.

[7] This is comparable to many other countries (World Bank 2007, Ch.9).

[8] As far as it was measured in the Sri Lanka data, enterprise income accounted for only 6% of the total, but this is definitely an understatement since 13.9% of the households had members working in an enterprise but not in wage employment (Table 1) and enterprise income was measured only if someone had a wage job. Given this bias, a decomposition of income in Sri Lanka is omitted in this section.

together for mutual gain.[9] This paper studies what households do, not what individual members in a household do.

As a discrete outcome, the econometric model may be formulated in a probit-like fashion. For household *i* in community *j*, the latent variable y_{ji}^* that determines the entrepreneurship outcome y_{ji} depends on household-level explanatory variables X_{ji} and community-level variables Z_j:

$$y_{ji}^* = X_{ji}'\beta + Z_j\gamma + \varepsilon_{ji} \tag{1}$$

$$\begin{aligned} y_{ji} &= 1 \quad \text{iff} \quad y_{ji}^* \geq 0 \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{2}$$

The disturbance ε_{ji} is likely to contain a community-level component that captures aspects of the investment climate that are not covered by observable factors Z_j. Thus,

$$\varepsilon_{ji} = \mu_j + \nu_{ji} \tag{3}$$

In this way, equations (1)-(3) constitute a typical random effects probit model.

Two complications present themselves. The first concerns sampling weights w_{ji}. Because of stratification by entrepreneurship status and clustering by community, accounting for w_{ji} is crucial for the development of the econometric model.[10] In an ordinary probit model, sampling weights are easily integrated into the likelihood function. In the absence of a community random effect, let $l_{ji} = l\left(\theta; y_{ji}, X_{ji}, Z_j\right)$ be the contribution of household *ji* to the likelihood function. The weighted log-likelihood function $\ln L = \sum_{j=1}^{J}\sum_{i=1}^{N_j} w_{ji} \ln l_{ji}$ implies a pseudo-likelihood function $L = \prod_{j=1}^{J}\prod_{i=1}^{N_j} l_{ji}^{w_{ji}}$, where it is understood that the weights average to 1. If we interpret each l_{ji} as a probability, the likelihood function with weights is therefore a joint probability that is a geometrically weighted combination of individual probabilities. In a random effects probit model, the combined contribution of all households *ji* in community *j* to the likelihood function is $l_j = \int\left(\prod_{i=1}^{N_j} l_{ji}\right) g\left(\mu_j\right) d\mu_j$ with $l_{ji} = l\left(\theta; y_{ji}, X_{ji}, Z_j, \mu_j\right)$: it is the product of the individual contributions, averaged over all feasible values of μ_j and thus a joint probability over all these households. Should sampling weights now be inserted in geometric fashion? Since household weights within a community are unlikely to sum to the number of

[9] E.g., Chiappori (1992), Browning and Chiappori (1998), Vermeulen (2002). In particular, models that address non-participation and household production are relevant; see, for example, Chiappori (1997), Donni (2003, 2008), and Blundell et al. (2007). For a different modeling approach to address this same issue, see Newman and Gertler (1994).

[10] Research on modeling under choice-based stratification goes back several decades. Related to discrete choice models, important contributions include Manski and Lerman (1977), Manski and McFadden 1981), Coslett (1981), and Wooldridge (2001). For a useful extension, see also Bhattacharya (2005). These econometric models typically assume that observations are independent. Bhattacharya (2005) assumes that observations may be correlated within a cluster, but his criterion function does not make this explicit.

households, i.e., $\sum_{i=1}^{N_j} w_{ji} \neq N_j$ in general, raising l_{ji} to the power of w_{ji} does not yield an expression for l_j that may be interpreted as a geometrically weighted joint probability at the community level. Instead, define the average household weight in community j as $\bar{w}_j = \sum_{i=1}^{N_j} w_{ji} \Big/ N_j$ and a community-standardized household weight as $\tilde{w}_{ji} = w_{ji} / \bar{w}_j$. The weighted pseudo-log-likelihood function is then defined as $\ln L = \sum_{j=1}^{J} \bar{w}_j \ln l_j$ with $l_j = \int \left(\prod_{i=1}^{N_j} l_{ji}^{\tilde{w}_{ji}} \right) g\left(\mu_j\right) d\mu_j$. In the latter expression, the integral may be simulated as the average of R random draws of μ_j^r from the distribution that is characterized by g:

$$\int \left(\prod_{i=1}^{N_j} l\left(\theta \mid y_{ji}, X_{ji}, Z_j, \mu_j\right)^{\tilde{w}_{ji}} \right) g\left(\mu_j\right) d\mu_j \approx \frac{1}{R} \sum_{r=1}^{R} \prod_{i=1}^{N_j} l\left(\theta \mid y_{ji}, X_{ji}, Z_j, \mu_j^r\right)^{\tilde{w}_{ji}}$$

This is an application of the GHK simulation estimation method (Hajivassiliou and Ruud, 1994; Vijverberg 1997).

The use of sampling weights also has consequences for the computation of the standard errors of the parameter estimates. As it is in the unweighted case, the variance of $\hat{\theta}$ is found by a Taylor expansion of the gradient of the pseudo-likelihood function:

$$\frac{\partial \ln L\left(\hat{\theta}\right)}{\partial \theta} = 0 = \frac{\partial \ln L\left(\theta\right)}{\partial \theta} + \frac{\partial^2 \ln L\left(\theta\right)}{\partial \theta \, \partial \theta'} \left(\hat{\theta} - \theta\right) \tag{4}$$

Thus

$$\begin{aligned} Var\left(\hat{\theta}\right) &= E\left[\left(\hat{\theta} - \theta\right)\left(\hat{\theta} - \theta\right)' \right] \\ &= \left(\frac{\partial^2 \ln L\left(\theta\right)}{\partial \theta \, \partial \theta'} \right)^{-1} \left(\frac{\partial \ln L\left(\theta\right)}{\partial \theta} \frac{\partial \ln L\left(\theta\right)}{\partial \theta'} \right) \left(\frac{\partial^2 \ln L\left(\theta\right)}{\partial \theta \, \partial \theta'} \right)^{-1} \end{aligned} \tag{5}$$

Unlike the unweighted case, the expected value of the outer product of the gradient is not equal to the expected value of the negative of the second order derivative, as is clear from the presence of weights w_{ji} and w_{ji}^2 in the following expressions (ignoring random effects for simplicity):

$$V_1 \equiv \left(-\sum_{j=1}^{J} \sum_{i=1}^{N_j} w_{ji} \frac{\partial^2 \ln l_{ji}\left(\hat{\theta}\right)}{\partial \theta \, \partial \theta'} \right)^{-1} \tag{6}$$

$$V_2 \equiv \left(\sum_{j=1}^{J} \sum_{i=1}^{N_j} w_{ji}^2 \left(\frac{\partial \ln l_{ji}\left(\hat{\theta}\right)}{\partial \theta} \right) \left(\frac{\partial \ln l_{ji}\left(\hat{\theta}\right)}{\partial \theta} \right)' \right)^{-1} \tag{7}$$

Thus, neither can serve as the variance of $\hat{\theta}$. Instead, the estimated variance is obtained by evaluating the first and second order derivatives at $\hat{\theta}$:

$$EstVar\left(\hat{\theta}\right) = \left(\frac{\partial^2 \ln L\left(\hat{\theta}\right)}{\partial\theta\,\partial\theta'}\right)^{-1}\left(\frac{\partial \ln L\left(\hat{\theta}\right)}{\partial\theta}\frac{\partial \ln L\left(\hat{\theta}\right)}{\partial\theta'}\right)\left(\frac{\partial^2 \ln L\left(\hat{\theta}\right)}{\partial\theta\,\partial\theta'}\right)^{-1} = V_1 V_2^{-1} V_1 \equiv V_3 \quad (8)$$

This expression is not sensitive to the scaling of the weights.

The second complication concerns the potential endogeneity of explanatory variables. A household-level explanatory variable X_{ji} may itself be driven by a random effects model where its household disturbance v_{ji}^x may be correlated with v_{ji} and its community factor μ_j^x with μ_j. At the community level, the disturbance μ_j^z driving a regressor Z_j may be correlated with μ_j. A Hausman test inserts estimated values $\hat{v}_{ji}^x$, $\hat{\mu}_j^x$ and $\hat{\mu}_j^z$ into the random effects probit model and examines their individual and joint significance. As it turns out, for the variables that were examined in this manner, these factors are never significant.[11]

4. RURAL ENTREPRENEURSHIP AND THE INVESTMENT CLIMATE

Definition of the Variables

In principle, household entrepreneurship is indicated by the operation of an enterprise by members of the household. Ideally, this would be directly observed as part of the survey effort, but the enterprise is not always located at the residence. Especially in Sri Lanka, standalone enterprises are common.[12] Two good alternatives are: (*i*) deriving a positive income from an enterprise, and (*ii*) household members reporting to be working on own account. These definitions should largely overlap, leaving a gap only if work on own account earned no income. In practice, some households indeed report work but no income; however, some others report income but no work. In Nicaragua, there is a discrepancy between the two definitions in 7.3 percent of the sample households, split evenly between those that report work but no income and those that report income but no work. In Sri Lanka, 1.3 percent reported income but no work, and, to a large degree because of the questionnaire design flaw, 29.6 percent reported work but no income.[13] The income-based definition is preferable from a perspective of productivity- and

[11] Given that the entrepreneurship equation is estimated with maximum likelihood, it makes sense to also pursue maximum likelihood methods to estimate the nature of the correlation among these potentially endogenous variables. The likelihood function of the entrepreneurship equation and the equations for one household-level variable and one community-level variable may be split into two components that separate the community-level equation from the rest. Unfortunately, the (conditional) likelihood function associated with the entrepreneurship and household-level variables is difficult to maximize and has not yet yielded usable estimates.

[12] Due to a design feature of the early RIC surveys, the household that owned and operated a standalone enterprise was not interviewed, and the subsample that did not operate a residence-based enterprise could—and did—report income from nonfarm enterprises. Thus, the rate of entrepreneurship based on the enterprise/household link in the RICS data reflects only household-based entrepreneurship and therefore understates the overall rate of rural entrepreneurship. Examining income receipt and work activity is a way around this design flaw.

[13] Note that these are unweighted counts. These consistency problems are quite common in survey data. In the RICS of Tanzania, 16.7 percent of the households received enterprise income but reported no work on own account, and 10.2 percent had household members working in a household enterprise but reported no income from a non-

poverty analysis—and is therefore used in the Nicaragua analysis—, but for Sri Lanka the work-based definition is an acceptable alternative.

The explanatory variables are divided into three sets. The first pertains to characteristics of the household and includes demographics such as the number of male and female adult members of the household, their age and human capital, and the gender and ethnicity of the head. These variables represent determinants of the available supply of household labor that may be drawn upon in a household enterprise, the quality of this labor, and barriers both to the operation of an enterprise and to other opportunities of employment. For example, these variables capture any barriers on the basis of gender, ethnicity or age, whether encountered when seeking wage employment in the labor market or in procuring inputs or financial resources for the enterprise. Similarly, human capital may signal skills used either in the enterprise (directly for production or indirectly for procuring inputs) or elsewhere in wage employment.

The set of household characteristics also includes a variable describing whether the household head's parents were entrepreneurs. The household head would have been able to experience entrepreneurship from up close, receiving an in-house apprenticeship as it were. Parents may have introduced the head to their business network.

The last group of household variables describes financial resources: household assets, remittance income, and income from other sources. First, assets could facilitate the operation of an enterprise, but they might be endogenously related to the decision to operate an enterprise. To be sure, the assets do not include enterprise assets that are available in the enterprise database, but it is possible that an entrepreneur uses household assets for his business, especially if the enterprise is based at the residence. Moreover, previous profitability may well have generated savings or investments in consumer durables that are now observed as household assets. The econometric analysis recognizes household assets as a potential endogenous household-level variable. Second, income from remittances and other sources may provide a cash flow that facilitates the ongoing operation of the enterprise—or it may relieve the need to operate an enterprise in order to earn an income. In Sri Lanka the measurement of remittances and other income is conditional on the household reporting any wage earnings due to a questionnaire flaw. Since this biases the measured values in ways that biases the parameter estimates, the income variables are omitted from the Sri Lanka analysis.

The second set of explanatory variables describes the investment climate in the community where the household resides. The investment climate is measured by a multitude of variables that are condensed into six benchmark indicators (Appendix A).[14] Both these indicators and their components are used in the econometric analysis in order to measure the impact of the investment climate on entrepreneurship. The benchmark indicators cover the characteristics of (*i*) connectivity with other regions, (*ii*) availability of infrastructure, utilities, and public services, (*iii*) availability of private business services, (*iv*) governance, (*v*) human capital, and (*vi*) finance services. These variables are constructed such that a higher value is expected to favor the performance of an enterprise. Connectivity may raise the value of entrepreneurship if it opens up distant markets or may lower it if distant competitors seek out

---

farm enterprise. For similar evidence in the context of Living Standards Survey data in Ghana, Guatemala, Kyrgyz Republic and Vietnam, where the percentage of inconsistencies ranged from 5 to 30 percent, see Vijverberg (2005).

[14] The literature that motivates these variables is enormous. Useful summaries are found in Haggblade et al. (2007), Lanjouw and Lanjouw (2001), Reardon (1997), and Reardon et al. (2001).

clients in the community. Infrastructure services are necessary for many business exploits: greater access should increase the value of entrepreneurship. The availability of business services may facilitate entrepreneurship. The governance index measures rule of law, security, contract enforcement, and so forth, without which the operation of an enterprise is more risky. The community human capital index is a proxy for the quality of labor that may raise the productivity of the enterprise and thus the value of entrepreneurship. At the same time, human capital is positively related to income and thus to potential market size. The finance services index summarizes finance and insurance services that facilitate the start-up and operation of an enterprise and thus should encourage household entrepreneurship.

However, as mentioned above, all of these arguments do not necessarily imply that the benchmark indicators must favor the entrepreneurship choice. A robust investment climate may spawn opportunities outside a household-owned-and-operated enterprise that are more lucrative and lead to enterprise closure—but also lead to an expanded non-farm economy nonetheless. The next section returns to this issue, but for now it is sufficient to temper expectations regarding the impact of the investment climate indicators on entrepreneurship choice.

The third set of explanatory variables describes the community apart from immediate associations with the policy-related investment climate. These variables are motivated as follows. Community size and the level of per capita income measure market opportunities that an enterprise could exploit inside a community. Since a successful enterprise in turn generates income, per capita income may be an endogenous community-level variable. The male wage rate measures the cost of labor, the opportunity cost of the entrepreneur and household members, and the level of well-being among potential customers: the impact on household entrepreneurship is therefore ambiguous. Enterprise openness indicates connection with markets outside the own community and thus, as it broadens the potential market, is expected to encourage entrepreneurship. Finally, in a community where agricultural seasonality is more pronounced, households may seek ways to make time in the slack season more profitable (Haggblade, Hazell and Brown, 1989; Lanjouw and Lanjouw, 2001); alternatively, the cyclical unreliable availability of labor may be a stumbling block for many enterprises.

Table A.1 of Appendix A provides definitions of the variables employed in this study; Table B.1 of Appendix B presents the descriptive statistics. The number of households with complete data is 1163 in Nicaragua and 849 in Sri Lanka.[15] The rate of entrepreneurship is 21.9% in Nicaragua and 25.3% in Sri Lanka. Wage employment is found in 70.4% and 76.0% of the households, respectively. Involvement in farming diverges between these countries: 27.2% in Nicaragua as opposed to 47.9% in Sri Lanka.

Regression Results

Entrepreneurship is a discrete 0/1 outcome and therefore is analyzed with a probit regression model that accounts for sampling weights as well as random effects by community.[16]

[15] In Nicaragua, 300 households were lost because their community of residence was not recorded in the household database. In Sri Lanka, community information needed to compute household weights or enterprise openness was missing for 162 households that were in communities in the north and east where civil unrest made survey logistics difficult. Other data losses were incidental.

[16] Deininger et al (2007) also estimate models for entrepreneurship selection and startup with Sri Lanka RICS data. Their econometric model does not account for community random effects and appears to omit sampling weights; investment climate variables are constructed in a different way; and for startup regressions households with an older enterprise are omitted as they are not considered to be part of the at-risk sample.

Two explanatory variables may be endogenous: household assets and the average income level in the community. In both countries, Hausman tests indicated no evidence that endogeneity might be an issue.[17]

The clustering of the households in communities makes it plausible that, aside from observable community characteristics, there may be unobservable community factors that are common to all households and cause correlation among the model's disturbance terms. This renders the standard probit model inadequate: a random effect must be incorporated. However, in cases where the estimated standard deviation of the random effect proved to be exceedingly small, the estimation strategy was simplified to a weighted probit model (without a random effect).

Table 2 reports the estimated models. For Nicaragua, the random effect always vanishes in estimation. For Sri Lanka, a random effect model is needed, since the standard deviation of community random effect is estimated at about 0.35, which is one third of the standard deviation of household-level random disturbance. Related to this, compared with Sri Lanka, the Nicaragua regressions find more community variables with statistically significant effects.

The first column of Table 2 contains only the six benchmark indicators; the second column adds the household characteristics; the third column fills in the remaining community variables. The fourth column selects variables from among the set of benchmark indicators, their components and other community variables that make the largest contribution to the explanation of entrepreneurship choice. Since parameter estimates of the probit model are tedious to interpret and the scales of the variables differ, the table shows the percentage point impact of a one-standard-deviation change in each explanatory variable, evaluated for an average household. For explanatory dummy variables such as "Head female", "Head parents entrepreneur", and "Not Sinhalese", the change in the explanatory variable is one unit, which is clearly more appropriate. In this, it is useful to keep in mind that the rate of entrepreneurship in Nicaragua is 22 percent. Thus, the second column indicates that a household with 0.99 additional male adults (which is the standard deviation of this variable – see Table B.1) is 3.08 percentage points more likely to operate an enterprise. Also, when parents of the household head were entrepreneurs, the

[17] When residuals of instrumental equations are added to columns 2-4 of Table 2, the highest *t*-statistic equals 1.39 in Sri Lanka, and the lowest p-value on a joint test equals 0.114 in Sri Lanka. As for Nicaragua, the estimated random effect has zero variance, such that only the household-level residual of the assets equation may be entered. Its t-statistics ranges from –1.72 in column 2 to –1.35 in column 4. Instrumental variables were selected from a larger group of candidate variables on the basis of their explanatory power. The household assets equation is estimated with a random effects model with sampling weights on the sample of households. The community income level equation is estimated by OLS on the sample of communities; the fact that this sample is a relatively small cross-section puts a limit on the number of variables that may be entered. Apart from variables that appear in the entrepreneurship equation, the instrumental variable household assets equation in Nicaragua contain the following variables: age squared; dummy for head able to read and write (having at least four years of schooling); and interactions between household human capital and (a) dummy for female head, (b) the community human capital and governance benchmarks. The community income level equation contains: distance to city and schools, and price of food. In the case of Sri Lanka, instrumental variables for household assets (apart from most variables already in the entrepreneurship equation) are age squared; dummy for head able to read and write (having at least four years of schooling); interactions between household human capital and (a) dummy for female head, (b) the access, community human capital, and finance benchmarks, and (c) enterprise openness index; and interactions between dummy for female head and the access, governance and finance benchmarks. Instrumental variables for community income level (in addition to most of the benchmarks) are distance indices to city, schools, and hospitals/clinics, price of rice, and acreage of upland and lowland per capita.

probability that the household operates an enterprise rises by 2.27 percentage points, although this effect is not estimated precisely as the *t*-statistic equals only 0.86 (see Table 2, part A).

In Nicaragua, entrepreneurship is more likely in households with more adults, whether male or female and in households with older members. Human capital does not matter; skilled workers may be more productive in the enterprise but, as all the literature on wage earnings shows, they are also able to receive higher wages when working for someone else. Female headed households are 6 percentage points more likely to operate an enterprise.

In Sri Lanka, where 25 percent of the households operate an enterprise, the effect of household size, age structure and human capital is smaller and statistically not significant. Female headship has a numerically strong negative effect, equal to 9 percentage points, but it, too, is not precisely enough estimated to attain statistical significance. Members of the ethnic (non-Sinhalese) minorities tend to not operate an enterprise, though its effect is also not precisely estimated. Surprisingly, the occupation of the head's parents proves to be irrelevant in both countries. Elsewhere, such as in Vietnam (e.g., Vijverberg and Haughton, 2004), the intergenerational effect is rather pronounced.

In regard to the household financial variables, the specification of the entrepreneurship model varies between the two countries, owing to the bias in the measured income components referred to above in Sri Lanka. In Nicaragua, access to remittance income and other income sources reduces the incentives for self-employment.

On the other hand, assets improve the chances. Specifically, the standardized effect of log-assets raises the probability of entrepreneurship by 5 percentage points in Nicaragua and 10 percentage points in Sri Lanka. To interpret this for Nicaragua where the average of log-assets amounts to US$ 1,419, a one-standard deviation change in log-assets equals US$ 6,485, which would take a household from, say the 25th percentile to about the 85th percentile of the distribution of assets. For Sri Lanka, the 10 percentage points increase in the likelihood of entrepreneurship comes about in response to a rise of US$ 11,621 in assets, relative to a base of US$ 5,405; this is equivalent to moving the household from the 25th percentile to the 80th percentile of the distribution of assets.[18] These computations show that household assets are an important determinant of entrepreneurship but also that establishing and operating an enterprise demands large sums of resources.[19]

Among the benchmark indicators, community human capital stands out in Nicaragua, where the standardized effect is about 5 percentage points. Infrastructure services has a positive effect of about 5.5 percentage points as well, but with the introduction of other community variables the effect diminishes to about 3. Business services may reduce the likelihood of entrepreneurship, and connectivity, governance and finance have no effect. The last column of Table 2, which contains only those community variables with the greatest explanatory power, lists positive effects of proximity to the post office and access to protected water: for both, the effect is 3 percentage points per standard deviation increase in the benchmark component.

[18] The distribution of assets is highly skewed. For both countries, the change in assets is also equivalent to the difference of the 75th and about 92nd percentiles.

[19] It is noted that this estimated effect is causal and not merely a reflection of reverse causation where prosperous entrepreneurs are accumulating household wealth: as mentioned above, assets were found to be exogenous relative to entrepreneurship choice.

The Sri Lanka results list a positive effect for connectivity, with the proximity to the post office highlighted in this country as well; the effect of 5 percentage points is quite large. The infrastructure services index, business services and governance and the human capital and finance indices do not appear to matter. The last column highlights sewage systems, which is found to have a strongly negative effect and thus contributes to the overall negative impact of the infrastructure services indicator; engineering services with a negative effect; and information technology services with a positive effect. The negative effects are counterintuitive, but these variables are all simple 0/1 indicators. As there are only 13 communities with a sewage system and merely 5 where engineering services are available, the estimation results may simply be the luck of the draw.

The last group of variables consists of various community characteristics. Here, Nicaragua and Sri Lanka yield opposite results. In Nicaragua, enterprise openness (i.e., selling and buying outside the community) raises the likelihood of entrepreneurship in a household by about 5 percentage points. The community income level may have a positive effect, but its effect is collinear with community human capital, the effect of which is more robust and therefore enters the last column. Agricultural seasonality, community size and the level of male wages do not matter. In Sri Lanka on the other hand, the likelihood of entrepreneurship is higher in larger communities with higher wages, although income per capita, seasonality and enterprise openness had no effect.

As mentioned before, Nicaragua's model is weighted probit without a random community effect. In Sri Lanka, the random community effect is statistically important. A one-standard deviation in the random effect moves the likelihood of entrepreneurship in a household up by almost 10 percentage point in the model with benchmark indicators and 7 percentage points in the model with selected benchmark components.

To conclude this examination of household entrepreneurship choice, it is worthwhile to reflect on the explanatory power of the model. The evidence is summarized in Table 3. The overall explanatory power of the model with household variables, benchmark indicators and community characteristics (Panel A) is 12.42 percent in Nicaragua and 7.85 percent in Sri Lanka. Household characteristics explain (i.e., improve the criterion function by) 8.09 percent in Nicaragua and 4.58 percent in Sri Lanka, where it should be noted that the Sri Lankan model does not include remittance and other income, which are influential variables in Nicaragua. The benchmark indicators help to explain 2.98 percent in Nicaragua and 1.46 percent in Sri Lanka, and community characteristics further add about half as much. In all, benchmark indicators and community characteristics contribute 35 percent of the explanation of entrepreneurship choice in Nicaragua and 28 percent in Sri Lanka. Further checks in Panels B, C and D confirm that benchmark indicators carry relatively less information in Sri Lanka than in Nicaragua.

5. ACTIVITIES OF HOUSEHOLDS

Nonfarm entrepreneurship is just one of several activities that people may select in order to earn a living. A study of entrepreneurship choice alone paints an incomplete picture. Therefore, this section extends the analysis of entrepreneurship choice to the choices that households make in regard to wage employment and farming. Moreover, as mentioned above, since during the course of economic development people may switch from self-employed farming into rural non-farm self-employment and wage employment, the percentage of

households that move into the nonfarm economy is of interest: it integrates the development of the non-farm economy into a single concept.

Estimation results are reported in Table 4, showing only the specifications with selected components. Each equation is estimated separately with a random effect probit model.[20] The first column refers to engagement in non-farm enterprises and is therefore identical to the fourth column of Table 2.

In Nicaragua, households with more men diversify into more activities; those with more women merely pursue non-farm entrepreneurship. In Sri Lanka, diversification is stronger yet when there are more adult males, and specialization is stronger when there are more adult women. The age pattern is similar in the two countries: households with older members are more likely to farm and perhaps to operate an enterprise; younger households seek wage employment. In Nicaragua they also engage more in entrepreneurship. Human capital draws households into wage employment, perhaps out of non-farm entrepreneurship, and likely out of farming. However, the entrepreneurial experience of the head's parents leaves no sorting imprint on the head's household.

Nicaragua and Sri Lanka differ in the sorting patterns according to the household head's gender. Female-headed households in Nicaragua are more likely to operate an enterprise, whereas those in Sri Lanka are significantly more likely found in wage jobs, rather than in a household enterprise. In both countries, however, farming is less likely for them, although the effect is large but imprecisely estimated. In Sri Lanka, non-Sinhalese are less likely to participate in the non-farm economy. In particular, they are found in farming rather than a non-farm business.

Extraneous income sources reduce the incentive for households to participate not only in a business, as seen above, but actually in any kind of work. This is consistent with the notion that, in standard labor economics terms, leisure is a normal good. Assets strongly sort households into both enterprises and farming; households with few assets seek wage employment.

The third column of Table 4 focuses on non-farm economic activity, combining (both farm and non-farm) wage employment and non-farm self-employment. For some explanatory variables, the third column sums contradictory signals from the first and second columns of the table; e.g., see the effect of age, female headship, ethnicity, or asset ownership. Among the strongest results, human capital pulls people into non-farm economic activity. In Sri Lanka, this is supplemented by a positive effect for adult women and negative effects for age and non-Sinhalese ethnicity. But balancing these insights is an equally valid inference that an exclusive study of non-farm economic activity hides interesting sorting patterns in the economy among wage and self-employment.

Searching for the best explanation through community variables, benchmark indicators and benchmark components yields various results. In Nicaraguan communities with higher incomes, households are more likely to engage in wage employment and less likely in self-employed farming, and when seasonality is more pronounced, farming is more likely and wage

[20] The choices are probably interdependent and some efficiency gain might be derived from joint estimation. A multinomial logit model incorporates all alternatives simultaneously but requires that individuals or households make exclusive choices. As seen in Table 1, households frequently participate in several activities, contrary to the logic of a multinomial logit application.

employment less so. Wage employment diminishes when more households have access to electricity or when government agencies are more transparent in their dealing with entrepreneurs in the community: these effects are counterintuitive, since one would expect the demand for wage labor to flourish under those conditions.[21] The weaker negative effect of human capital in the community should be seen in the light of the positive private effect of human capital: one's comparative advantage in skills disappears when other members in the community are skilled as well. For farming, the data suggest that the more rural and remote the community, the more likely will households earn their living by farming, although the estimates are not precise enough to imply statistical significance. Finally, the community random effect may be absent in the determination of entrepreneurship choice, but it is highly significant in households' choice of wage employment and farming: the standardized effect ranges from 15 to 19 percentage points.

In Sri Lanka, the list of community variables and benchmark components is long and varied. Households in larger communities are more likely to work for a wage and less likely to farm. Greater seasonality is associated with farming choice. Turning to the benchmark components, it appears that chances of wage employment are higher in communities farther from a major city or closer to a main market. Overall access to infrastructure helps; more detailed specifications show positive effects for the percentage of households with electricity and the presence of concrete or asphalt roads and a negative estimate for the percentage of households with cell phones. Furthermore, according to the estimates, wage employment is more likely with the presence of marketing services available to businesses in the community and less likely with higher "financial penetration," i.e., the percentage of entrepreneurs in the community that have considered asking a financial institution for a loan in years prior to the RIC survey. Farming is more prevalent in communities with poorer access to electricity or phone service. As the last variable, financial penetration has a positive effect, as farming may also benefit from access to credit.[22] Across all outcome variables in Table 4B, the random community effect is influential. In particular, the prevalence of farming varies greatly from community to community in ways that are not captured in the observable information about communities. This is consistent with the notion that farming in a community depends more than other enterprises on land and water availability in the location than on investment climate variables.

6. *Enterprise Start-Up*

Measuring Start-Up

The presumption examined in this paper is that the investment climate sets the tone for entrepreneurial activity. Entrepreneurship is a dynamic activity, one aspect of which is reflected in enterprise start-up.

Unfortunately, the RICS survey does not measure start-up directly. Rather, a start-up evidently occurred when a household is operating a "new enterprise"—in this study, defined as one that started its operation during the current or previous calendar year. This proxy is not perfect. For one, it understates the true two-year start-up rate, since these enterprises must have

[21] These conditions might provide incentives to start one's own business instead, but column 1 of the table reveals no such effect.

[22] The effect of management consulting and marketing services is strong and statistically significant and in a plausible direction—but may have to be taken with a grain of salt, as the number of communities with such services (9 and 8 respectively, with an overlap of 6 between them, out of 117) is small.

survived for up to two years to be included: enterprises that recently started up but already died before the survey was held cannot be counted. Furthermore, because of the design of the survey, such a start-up enterprise is by necessity based on the premises of the household's residence at the time of the survey. Thus, enterprises that started up as or converted into standalone units are not counted either.

First, consider the case of Nicaragua. The enterprise start-up rate is measured to be 3.0 percent (Table A.1).[23] That is, 3% of the Nicaraguan households started a household-based enterprise during the two years prior to the RICS survey. There were actually only a few households with inconsistent enterprise income and work activity in the database. Thus, only a few of the standalone enterprises might be under household ownership of one of the households in the sample, and thus any correction for start-up among standalone enterprises[24] would be minor. However, because the Nicaragua RICS was administered early (February-May) in the year, the 3.0% figure still understates the true two-year start-up rate. It is safe to assume that the Nicaragua survey may be capturing only one third to one half of start-ups in 2004. Thus, the estimated Nicaragua start-up rate should be revised upward to 4.0 to 4.5%.

In Sri Lanka, the situation is quite different. The rate of enterprise start-up is measured to be 1.30% (Table A.1). The weighted percentage of households with a (household-based) enterprise in the survey is 5.47%. Thus, 21.93% of these enterprises in the survey are rated as start-ups. But another 20.30% of the households report work on own account but did not provide data about the enterprise. Now in the RICS *enterprise* database, 35.24% of the enterprises are household-based and 64.76% are standalone enterprises; 24.96% and 23.90% of them are start-ups, respectively.[25] If one were to infer that the 20.30% of the households that report work on own account but did not provide data about the enterprise are operating a standalone enterprise and that the start-up rate of 23.90% applies to them uniformly, then another 4.85% (= 0.239 x 20.30) of the households would be operating a start-up enterprise, and the enterprise start-up rate among the Sri Lanka rural households would be 6.05%.[26] Again, this statistic further assumes that standalone start-ups did not first begin as a household-based entity that was turned into a standalone within the first two years of its existence. If this happens, double counting occurs among the start-up percentage.

One more feature matters for the comparison between Nicaragua and Sri Lanka. The Nicaragua RICS surveyed every enterprise activity in the household, which raises the chance that a start-up is observed in a household; indeed, the start-up rate among activities in Nicaragua is 13.28% but among households is 18.15% (see footnote 23). The Sri Lanka RICS selected only one main activity, which is likely the more established one. By virtue of these survey design features, the start-up rate in Sri Lanka should be relatively lower than Nicaragua.

---

[23] As for Nicaragua, there are 1372 household-based enterprises in 1060 households. Of these, 179 began operations during the previous two years (13.28%). These are aggregated to the household level, and a household is assumed to have started an enterprise when one of its several operations began during the previous two years; this means that 18.15% of the households with one of more enterprises began at least one. Averaging over all households with sample weights yields the measured 3.0%.

[24] Standalone enterprises were only 7.5% of all enterprises. The start-up rate among standalone enterprises was 21.3%.

[25] The difference between 24.96% and 21.93% arises from the fact that not all household-based enterprises are actually linked with a household and also that the second statistic is computed in a subsample that is subject to data availability.

[26] The Sri Lanka RICS was held late in the year (September-November), and 88 % of the existing enterprises started up in the months of January to August, as did most of the measured start-up enterprises. Thus, a revision because of the timing of the survey will not add much to the measured start-up rate.

There are two reasons that should lend credence to these estimated enterprise start-up rates. First, just as the start-up rate in Nicaragua is below that of Sri Lanka, the rate of entrepreneurship in Nicaragua is lower as well. Second, if it is assumed that the population of enterprises is stationary such that new start-ups merely replace exiting enterprises, enterprise start-up information may be translated into enterprise survival estimates. If so, the annual enterprise survival rate[27] among household-based and standalone enterprises equals 86.6 percent and 87.2 percent, respectively in Sri Lanka and 90.7 percent and 84.7 percent in Nicaragua, where in the latter case it is assumed that the data capture only half of the start-ups in 2004. By comparison, reported enterprise survival rates are about 83 percent in Vietnam (Vijverberg 2006), 66% in the Dominican Republic and 79% in Zimbabwe, Swaziland and Kenya (Liedholm 2007).[28]

Regression Results

Table 5 examines the determinants of enterprise start-up. These are estimates of weighted probit models, expressing the effect of explanatory variables again in standardized form. The community random effect vanished in both countries as its estimated standard deviation converged to 0. Moreover, because the number of start-ups is so small, the model could not be populated with a full set of benchmarks and community characteristics. This very fact implies that the estimates of the effect of community variables should be cautiously interpreted.

Actually, few explanatory variables matter and the estimated effects are small. This is understandable: the average start-up rate is only 3.0 percent in Nicaragua and 1.2 percent in Sri Lanka.[29] Among the statistically significant estimates, in Nicaragua only households receiving income from remittances and other sources are less likely to start up a new enterprise. In Sri Lanka, assets have a positive effect; the number of female adults and the entrepreneurial experience of the head's parents may raise the likelihood of start-up; and chances are lower among households that are older or are of non-Sinhalese ethnicity.

Among benchmark indicators, connectivity hurts chances in Nicaragua but has no impact in Sri Lanka. Proximity to a major city has a negative effect in Nicaragua but a positive one in Sri Lanka. Infrastructure services have a positive effect in Nicaragua but a negative impact in Sri Lanka. Community human capital raises chances in Sri Lanka. The remaining indicators do not matter.

Other community characteristics show a similar disagreement between the two countries. Enterprise start-up is more likely in Nicaragua in larger communities where enterprises are used to trading with clients outside the community and where agricultural seasonality is more pronounced. Of these three variables in Sri Lanka, enterprise openness as an opposite effect and the other two do not matter.

---

[27] If p measures the annual proportion of enterprises that exits and is replaced by new start-ups, the two-year ratio of start-up enterprises relative to the total number of enterprises equals $p(1-p)+p$, which then may be equated to the two-year percentage q of start-up enterprises in the sample. The survival rate equals $1-p$ and solves to $(1-q)^{0.5}$. The assumption of a stationary population of enterprises is crucial in identifying this survival rate.

[28] Using data from four African countries, McPherson (1995) reports firm closure rates of around 3%. His analysis is based on recall data on closed enterprises, which in other contexts has proven to understate enterprise closure.

[29] By comparison, in Table 2 where the estimated effects are larger, the average rate of entrepreneurship was 22 percent (Nicaragua) and 25 percent (Sri Lanka).

Overall, therefore, there is much variation in the impact of the investment climate between the two countries. It is possible that behavior differs; as mentioned earlier in this chapter, the impact of household and community variables is a priori ambiguous. Alternatively, these may simply be random patterns that appear when communities are few and variation in the dependent variable is limited. Moreover, one should expect differences between these two countries since the proportion of standalone entities is much greater in Sri Lanka: the analysis concerns start-up of household-based enterprises only, omitting start-up of standalone enterprises.

7. CONCLUDING REMARKS

This paper looks at only one facet of entrepreneurship choice: whether to set up and operate a business. This choice must be placed in context. If farming is more profitable than anything else, rational people will operate a farm. If a wage job pays better, many people will elect to work for an employer. If one's own enterprise is the most profitable option, it makes sense to start a business. Moreover, income may not be the only consideration in this comparison, as each option has its own non-monetary compensation and imposes its own income risk. Households may also pursue a diversification strategy.

Apart from household characteristics, a long list of features of the economic environment—the investment climate—may impact this choice. This is not to say that all of them will, at least to a measurable degree. The econometric analysis merely finds out what the most important factors are. Moreover, it is easy to argue that many investment climate factors have an ambiguous effect on entrepreneurship: they may raise productivity at both small enterprises (run by a household) and large enterprises (run perhaps by a corporate entity or a successful entrepreneur). Thus, an entrepreneur who sees his enterprise flourish may nevertheless abandon it when working in service of someone else becomes more profitable. This suggests that a parallel analysis should focus on enterprise productivity, distinguishing between large and small enterprises if possible.

According to the estimates, community characteristics matter in different ways in Nicaragua and in Sri Lanka. Entrepreneurship rises with community human capital, access to infrastructure, and enterprise openness in Nicaragua, and with connectivity (in particular, proximity to post office) and community wages in Sri Lanka. In Nicaragua, the most important elements of the investment climate determining involvement in wage employment and farming are the income level in the community and the degree of seasonality. In Sri Lanka, farming is more likely in smaller communities with more seasonality, poorer phone connection and greater financial penetration; wage employment is more common in larger communities closer to markets (but farther from a major city) with less financial penetration and better infrastructure services. The point is: community characteristics help in the labor market sorting process but are not always unambiguous in this. The investment climate generates a rich sorting pattern across the different economic activities, and a focus on non-farm economic activity as an aggregate can lead to an inadequate insight into the process of development.[30]

[30] The RICS data examined here do not distinguish wage employment by sector. In future administrations of this or similar surveys, it would be useful to separate agricultural and non-agricultural laborers.

This study has limitations. Activity choice is modeled with three separate regression equations, each estimated independently. This actually hides their interdependence. Greater efficiency is feasible by simultaneously estimating the three equations, but this comes at a cost of complexity. The structure of the data has households clustered in communities. In consequence, the model must allow for unobservable community factors (i.e. random effects), which, in the single equation approach pursued here, proved to be important. Estimation of a single random effect probit equation is already a more laborious procedure than the simple, common probit model; simultaneously estimating three equations will be time consuming.[31]

Another limitation is imposed by the survey design. By its very nature, the investment climate is a community feature that consists of a myriad of elements, but the survey covers only 100 to 150 communities for budgetary and logistical reasons. In fact, the households in the research samples reside in 93 Nicaraguan communities and 117 communities in Sri Lanka. Community-level variables can explain only the between-community variation in the dependent variable. This means that the investigation of the impact of the investment climate works with roughly 100 data points. Yet, the investment climate has many dimensions. This makes reliable detection of an effect of the investment climate a challenge.

To get around this problem, community variables may be aggregated into "benchmark indicators" as they were referred to in this paper in order to capture as many facets of the investment climate as possible—under the notion that the many community variables represent a smaller number of latent investment climate factors. Indeed, the six benchmark indicators alone consist of some 30 components. But the aggregative nature of benchmarks may also be a weakness. If only one component matters, its estimated effect is weighed down by its pairing with a number of other ineffective components. It is therefore advisable to still search among benchmark components for variables that provide strong explanatory power. Yet, there are so many components, and each is listed as a candidate variable because it reflects an aspect of the investment climate and may matter on economic grounds for entrepreneurship. The eventual selection is the result of, essentially, a statistical rather than economic comparison, i.e., a stepwise regression procedure. The selection is then vulnerable to accidental statistical significance.[32]

Apart from what is discussed in this paper, there are other choices that entrepreneurs make: which sector to enter, how to arrange for financing, which technology to use, where to locate, and how to seek out clients, to mention a few. In this sense, the scope of this paper is limited. However, RICS data offer scope for further analysis to better understand the full dimension of entrepreneurship choice.

The econometric model accounts for potential endogeneity of explanatory variables, such as income levels in the community and assets in the household. The analysis above failed to disprove the exogeneity of these variables, but more research is needed to successfully incorporate endogenous explanatory variables. Apart from this, it is quite conceivable that

[31] An alternative approach to the simple 0/1 modeling strategy would be to describe how many household members and even how much of their time is devoted to each economic activity. This converts the entrepreneurship decision into a continuous variable with a lower bound at 0. It also extracts more information from the survey data that is in fact available in the RICS data. Yet one step further is to focus on each member's activity choice; e.g., see Newman and Gertler (1994).

[32] Nevertheless, experimentation has shown that a few well-selected components and community characteristics can provide equal or better explanatory power than a full list of benchmark indicators and community characteristics.

remittance income—and the migration that precedes it—actually reflects activity choices of the household more so than act as a driver of the activity choice. A household may elect to send one of its members to a labor market elsewhere instead of employ this person in its own local enterprise. In this sense, remittance income would be an inappropriate determinant of household entrepreneurship, just as wage earnings or income from farming are inappropriate. This is another topic for future study.

8. REFERENCES

Batra, G., D. Kaufmann, and A. H. W. Stone, 2003. "Investment Climate Around the World: Voices of the Firms from the World Business Environment Survey." Washington DC: IBRD.

Bigsten, A. and M. Soderbom, 2006. "What Have We Learned from a Decade of Manufacturing Enterprise Surveys in Africa?" World Bank Research Observer, 21:2, 241-65.

Bhattacharya, Debopam, 2005. "Asymptotic inference from multi-stage samples." Journal of Econometrics, May, 126:1, 145-171.

Bils, Mark, and Peter J. Klenow, 2000. "Does schooling cause growth?" American Economic Review, December, 90:5, 1160-1183.

Blundell, Richard, Pierre-Andre Chiappori, T Magnac, and Costas Meghir, 2007. "Collective labor supply: heterogeneity and non-participation." Review of Economic Studies, April, 74:2, 417-445.

Browning, Martin, and Pierre-Andre Chiappori, 1998. "Efficient intra-household allocation: a general characterization and empirical tests." Econometrica, November, 66:6, 1241-1278.

Chiappori, Pierre-Andre, 1992. "Collective labor supply and welfare." Journal of Political Economy, June, 56:1, 63-89.

Chiappori, Pierre-Andre, 1997. "Introducing household production in collective models of labor supply." Journal of Political Economy, February, 105:1, 191–209.

Coslett, Stephen R., 1981. "Efficient estimation of discrete-choice models." In C.F. Manski and D. McFadden, eds., Structural Analysis of Discrete Data with Econometric Applications. Cambridge, MA: MIT Press, Ch.2.

Deininger, Klaus, Songqing Jin, and Mona Sur, 2007. "Sri Lanka's rural non-farm economy: Removing constraints to pro-poor growth." World Development, 35:12, 2056-2078

Dollar, David, Mary Hallmark-Driemeyer, and Taye Mengistae, 2005. "Investment climate and firm performance in developing economies." Economic Development and Cultural Change, October, 54:1, 1-31.

Donni, Olivier, 2003. "Collective household labor supply: non-participation and income taxation." Journal of Public Economics, May, 87:5-6, 1179–1198.

Donni, Olivier, 2008. "Labor supply, home production, and welfare comparison." Journal of Public Economics, July, 92:7, 1720-1737.

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

Haggblade, Steven, Peter B. R. Hazell, and Thomas Reardon, eds. 2007. Transforming the Rural Nonfarm Economy: Opportunities and Threats in the Developing World, Baltimore: The Johns Hopkins University Press for the International Food Policy Research Institute.

Hajivassiliou, Vassilis A., and Paul A. Ruud, 1994. "Classical estimation methods for LDV models using simulation." In R.F. Engle and D.L. McFadden, eds., Handbook of Econometrics, Vol 4, 2382-2441. Amsterdam (New York): Elsevier.

Hayami, Yujiro, ed., 1998. Towards a Rural-Based Development of Commerce and Industry, Selected Experiences from East Asia. Washington DC: Economic Development Institute of the World Bank.

Lanjouw, Peter, 2001. "Nonfarm employment and poverty in rural El Salvador." World Development, March, 29:3, 529-547.

Lanjouw, Jean O., and Peter Lanjouw, 2001. "The rural non-farm sector: issues and evidence from developing countries." Agricultural Economics, 26:1, 1-23.

Liedholm, Carl, 2007. "Enterprise dynamics in the rural nonfarm economy." In S. Haggblade, P.B.R. Hazell, and T. Reardon, eds. 2007. Transforming the Rural Nonfarm Economy: Opportunities and Threats in the Developing World, Baltimore: The Johns Hopkins University Press for the International Food Policy Research Institute, Ch.5.

Manski, Charles F., and Steven R. Lerman, 1977. "The estimation of choice probabilities from choice based samples." Econometrica, November, 45:8, 1977-1988.

Manski, Charles, and Daniel McFadden, 1981. "Alternative estimators and sample designs for discrete choice analysis." In C.F. Manski and D. McFadden, eds., Structural Analysis of Discrete Data with Econometric Applications. Cambridge, MA: MIT Press, Ch.1.

McPherson, Michael A. 1995. "The hazards of small firms in southern Africa." Journal of Development Studies. October, 32:1, 31-54.

Moock, Peter, Philip Musgrove, and Morton Stelcner, 1990. "Education and earnings in Peru's informal nonfarm family enterprises." Living Standard Measurement Study Working Paper 64, Washington DC: World Bank.

Newman, John L., and Paul J. Gertler, 1994. "Family productivity, labor supply and welfare in a low income country." Journal of Human Resources, Fall, 29:4, 989-1026.

Reardon, Thomas, 1997. "Using evidence of household income diversification to inform study of the rural nonfarm labor market in Africa." World Development, May, 25:5, 735-747.

Reardon, T., J. A. Berdegué and G. Escobar, 2001. "Rural Nonfarm Employment and Incomes in Latin America: Overview and Policy Implications." World Development, March, 29:3, 395-409.

Van der Sluis, Justin, Mirjam van Praag, and Wim Vijverberg, 2005. "Entrepreneurship selection and performance: A meta-analysis of the impact of education in developing countries," World Bank Economic Review, November, 19:2, 225-261.

Vermeulen, Frederic, 2002. "Collective household models: principles and main results." Journal of Economic Surveys, September, 16:4, 533-564.

Vijverberg, Wim P.M., 1990. "Non-farm self-employment and the informal sector in Côte d'Ivoire: a test of categorical identity." Journal of Developing Areas, July, 24:4, 523-542.

Vijverberg, Wim P.M., 1997. "Monte Carlo evaluation of multivariate normal probabilities." Journal of Econometrics, January, 76:1-2, 281-307.

Vijverberg, Wim P.M., 1998. "Nonfarm household enterprises in Vietnam." In David Dollar, Paul Glewwe, and Jennie Litvack, eds., Household Welfare and Vietnam's Transition, Washington DC: World Bank, Chapter 5, 137-178.

Vijverberg, Wim P.M., 1999. "The Impact of Schooling and Cognitive Skills on Income from Non-Farm Self-Employment," in Paul Glewwe, ed., The Economics of School Quality Investments in Developing Countries: An Empirical Study of Ghana, New York: St. Martin's Press, Chapter 6, 206-252.

Vijverberg, Wim P.M., 2005. "Rural investment climate and non-farm entrepreneurship in Living Standards Surveys." Manuscript, ARD, World Bank, February.

Vijverberg, Wim, with Hoang Thi Thanh Huong, Nguyen Chien Tang, Nguyen Ngoc Que, Nguyen The Quan, Phung Duc Tung, and Vu Thi Kim Mao, 2006. "Non-farm household enterprises in Vietnam: A research project using data from VHLSS 2004, VHLSS 2002 and AHBS 2003." The World Bank, Hanoi Office, July.

Wooldridge, Jeffrey M., 2001. "Asymptotic properties of weighted M-estimators for standard stratified samples." Econometric Theory, April, 17:2, 451-470.

World Bank 2004. World Development Report 2005: A Better Investment Climate for Everyone. New York: Oxford University Press and World Bank.

World Bank 2006. The Rural Investment Climate: It Differs and it Matters. Washington, DC: IBRD.

World Bank, 2007, World Development Report 2008: Agriculture for Development. Washington, DC: IBRD.

World Bank, 2008. The Rural Investment Climate: Analysis and Findings. Washington, DC: IBRD.

APPENDIX A: VARIABLE DEFINITIONS

Table A.1 provides definitions of the variables used in the analysis of this paper. The use of benchmark indicators as summaries of the investment climate in a community requires more detail. In general, benchmark indices are computed as the average of subindices, transforming the average value such that the maximum across all communities equals 1. Each subindex is constructed such that a lower quality of the investment climate feature is characterized by a smaller value of the subindex and is scaled to the interval from 0 to 1. Depending on its content, a subindex may be computed from the community database, the enterprise database or the household database. For more detail, see World Bank (2008, Annex E). The following description briefly lists the components that make up each benchmark.

1. Connectivity

The index of rural-urban connectivity/regional economic integration consists of eight subindices:

1) *Time taken by main means of transportation to the nearest major city.*
2) *Relative cost of transportation to the nearest major city (by public transportation).*
3) *Time taken by main means of transportation to the main market.*
4) *Cost of transportation to the main market (by bus).*
5) *Distance to the post office.*
6) *Rail stop in walking distance.*
7) *Percentage of households with fixed-line telephones.*
8) *Percentage of households with a cell phone.*

2. Infrastructure Services

The index that measures access to infrastructure services and service delivery consists of eight subindices:

1) *Percentage of households that use electricity.*
2) *Availability of electricity.*
3) *Percentage of households with access to drinking water from a protected source.*
4) *Percentage of households with a fixed-line telephone.*
5) *Percentage of households with a cell phone.*
6) *Sewage channels in the community.*
7) *Garbage collection or disposal service in the community.*
8) *Most common road surface (internal road) is concrete or asphalt.*

3. Business Services

The index that measures the availability of business services is derived from seven dummy variables that are measured in each community.

1) *Engineering services available for businesses in the community.*
2) *Management consulting services available for businesses in the community.*
3) *Marketing services available for businesses in the community.*
4) *Accounting services available for businesses in the community.*
5) *Legal services available for businesses in the community.*
6) *Insurance services available for businesses in the community.*
7) *Information services available for businesses in the community.*

4. Governance

This index measures existence of governance (or administrative burden) and corruption. It is the average of five subindexes that in turn summarize specific features of the investment climate:

1) *General policy and institutional constraints.*
   - **a.** Corruption as a constraint for rural investment climate;
   - **b.** Economic policy as a constraint for rural investment climate;
   - **c.** Crime, theft, and disorder as a constraint for rural investment climate;
   - **d.** Legal system as a constraint for rural investment climate; and
   - **e.** War or other social frictions as a constraint for rural investment climate.
2) *Infrastructure and services.*
   - **a.** Extra fee to register/renew;

b. Extra fee to apply for a basic activity license/permit/renew;
c. Extra fee to apply for a construction permit;
d. Extra fee to apply for electricity connection for industry use;
e. Extra fee to apply for electricity connection for domestic use; and
f. Extra fee to apply for other services.

3) *Dealing with government agencies.*
   a. Tax-related issues;
   b. Labor-related issues;
   c. Issues related to fire and building safety;
   d. Issues related to sanitation and epidemiology;
   e. Issues relating to environmental regulation; and
   f. Others.
4) *Rule of law.*
   a. Predictability of laws and regulations that affect the operation and growth of businesses; and
   b. Rules and regulations that can be manipulated or misinterpreted by officials.
5) *Conflict resolution and contract enforcement.*
   a. Must rely on the reputation of others with whom you enter into agreements;
   b. A contract will protect you from being cheated by others; and
   c. The legal system will uphold your contract and property rights in disputes.

5. Human Capital

This indicator provides an index of human capital constructed with information from the household survey. Bils and Klenow (2000) describe human capital stock according to the classic Mincerian returns to schooling and experience:

$$\ln h = \frac{\theta s^{1-\psi}}{1-\psi} + \gamma_1 (a - s - 6) + \gamma_2 (a - s - 6)^2$$

On the basis of Bils and Klenow (2000), the following parameter values are inserted: $\gamma_1 = 0.0512$, $\gamma_2 = -0.00071$, $\theta = 0.099$, and $\psi = 0$ (which corresponds to the classic case in the labor literature showing no diminishing returns to education). The human capital stock is computed for each individual between 16 and 65 years of age in the household sample and then averaged up to the community level and scaled.

6. Finance Services

This indicator describes the level of development of financial intermediation, measured with four subindices:

1) *Number of formal financial and insurance institutions that offer services in the community, weighted by distance from the community.*
2) *Number of financial and insurance services provided in the community by formal institutions, weighted by distance from the community.*
3) *Community access to a commodity exchange for futures or options contracts.*
4) *Access to loans.*

Table A.1: Definition and descriptive statistics of variables used in the econometric models

Variable	Definition
Household Characteristics	
Entrepreneurship[a]	Dummy variable, 1 = household operates an enterprise, 0 = otherwise
Wage employment[a]	Dummy variable, 1 = members of the household hold a wage job, 0 = otherwise
Non-farm economic activity[a,b]	Dummy variable, 1 = household operates an enterprise and/or household members hold a wage job, 0 = otherwise
Farming[a]	Dummy variable, 1 = household operates a farm, 0 = otherwise
Enterprise start-up[c]	Dummy variable, 1 = household operates an enterprise that started up at most two years prior to the time of the survey, 0 = otherwise
Number of male adults	Number of male adults in the household
Number of female adults	Number of female adults in the household
Average age	Average age among adults in the household
Human capital index	Average human capital[d] among adults in the household
Head female	Dummy variable, = 1 if the head of household is female, = 0 otherwise
Not Sinhalese	Dummy variable, = 1 if the head of household is not of Sinhalese ethnicity, 0 = otherwise
Head's parents were entrepreneur	Dummy variable, = 1 if the parents of the head of household, 0 = otherwise
ln(Other income)	Natural log of the sum of all types of income except wage, enterprise, farm and remittance income (USD)
ln(Remittances)	Natural log of remittance income (USD)
ln(Assets)	Natural log of household assets (USD)
Benchmarks and components	
Connectivity index	Index of rural/urban connectivity and regional economic integration[d]
Infrastructure services index	Access to infrastructure services and service delivery[d]
Business services index	Index measuring the availability of business services[d]
Governance index	Index measuring the existence of governance and corruption[d]
Human capital index	Community-level human capital index[d]
Finance services index	Index of development of financial intermediation[d]
Community variables	
ln(Community size)	Natural log of number of residents in the community
ln(Income per capita)	Natural log of the average income per capita (USD) in the community
Agricultural seasonality	Average of the standard deviation in monthly agricultural labor input for male and female labor
Enterprise openness	Index of dealings of the enterprise outside the community with respect to clients and input providers
ln(Male wage rate)	Natural log of the average male wage rate in agriculture, service, and manufacturing

Notes: [a] Dependent variable. Nicaragua follows income-based definitions; Sri Lanka follows work-based definitions.
[b] Combining non-farm entrepreneurship and wage employment.
[c] Dependent variable.
[d] Defined in the text, with more details provided in World Bank (2008, Annex E).

APPENDIX B: DESCRIPTIVE STATISTICS

Table B.1: Descriptive statistics of the Nicaragua and Sri Lanka RICS data

	Nicaragua		Sri Lanka	
Variable	Mean	StDev	Mean	StDev
Household Characteristics				
Entrepreneurship[a]	0.219	0.414	0.253	0.435
Wage employment[a]	0.704	0.456	0.760	0.427
Non-farm economic activity[a,b]	0.817	0.387	0.894	0.308
Farming[a]	0.272	0.445	0.479	0.500
Enterprise start-up[c]	0.030	0.171	0.013	0.112
Number of male adults	1.316	0.992	1.496	0.914
Number of female adults	1.459	0.901	1.494	0.783
Average age	33.975	8.726	37.323	7.701
Human capital index	0.209	0.136	0.333	0.158
Head female	0.329	0.470	0.191	0.393
Not Sinhalese	n.a	n.a	0.140	0.347
Head's parents were entrepreneur	0.362	0.480	0.272	0.445
ln(Other income)	1.190	2.345	n.a	n.a
ln(Remittances)	1.574	2.580	n.a	n.a
ln(Assets)	7.258	1.717	8.595	1.147
Benchmarks and components				
Connectivity index	0.350	0.186	0.436	0.166
Infrastructure services index	0.535	0.256	0.440	0.162
Business services index	0.296	0.428	0.179	0.241
Governance index	0.686	0.111	0.677	0.037
Human capital index	0.205	0.074	0.324	0.092
Finance services index	0.288	0.244	0.357	0.098
Community variables				
ln(Community size)	8.702	1.929	7.539	0.594
ln(Income per capita)	5.824	0.702	8.144	0.554
Agricultural seasonality	0.721	0.413	0.731	0.307
Enterprise openness	1.438	0.349	2.227	0.785
ln(Male wage rate)	0.002	0.199	5.542	0.200
Number of households	1163		849	

Notes: [a] Dependent variable. Column percentages do not add to 100 as households may participate in several activities simultaneously. Nicaragua follows income-based definitions; Sri Lanka follows work-based definitions.

[b] Combining non-farm entrepreneurship and wage employment.

[c] Dependent variable. In Sri Lanka, enterprise start-up is measured for 790 households, rather than 849.

[d] Defined in Appendix A.

Source: RICS data

Figure 1: Components of household income by quintile: Nicaragua

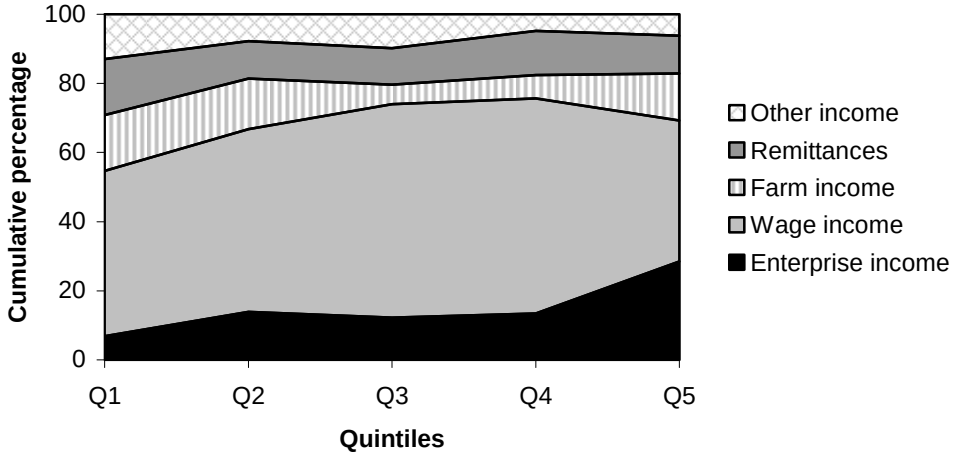

Table 1: Accounting for economic activities of households and household members

	Activity of the overall household, measured by			Employment[a] of adult household members		
	Employment[a] of adult household members	Any income[b] from stated source	Income from stated source > $60[c]	Total	Male	Female
Nicaragua						
Enterprise	7.2	7.7	8.7	10.1	8.5	11.5
Enterprise & Wage	9.4	8.3	8.9	2.5	2.7	2.3
Enterprise & Farming	2.6	3.5	2.7	1.6	2.3	0.9
Enterprise & Wage & Farming	3.5	2.7	1.7	0.5	0.7	0.3
Wage	37.3	45.5	50.5	29.2	36.4	23.1
Farming	9.3	10.0	8.5	13.0	18.1	8.6
Wage & Farming	19.7	12.8	7.4	10.0	18.3	2.9
Not employed	11.1	9.5	11.8	33.1	13.0	50.4
Total	100.0	100.0	100.0	100.0	100.0	100.0
Sri Lanka						
Enterprise only	8.6	0.0	1.5	8.8	10.2	7.4
Enterprise & Wage	9.1	9.9	10.8	1.5	2.7	0.2
Enterprise & Farming	5.3	1.5	0.3	2.8	3.2	2.4
Enterprise & Wage & Farming	6.4	6.6	4.3	0.4	0.7	0.0
Wage only	36.7	36.0	43.1	29.3	39.4	19.5
Farming only	8.4	18.3	13.3	8.9	9.9	7.9
Wage & Farming	22.5	26.4	20.1	5.8	10.2	1.5
Not employed	3.0	1.4	6.7	42.6	23.7	61.2
Total	100.0	100.0	100.0	100.0	100.0	100.0

Notes: [a] Employment refers to the previous 12 months.

[b] For Sri Lanka, adults living in 149 households with zero or negative total incomes are excluded. Moreover, income received from non-farm enterprises is understated because of a flaw in the household module of the questionnaire.

[c] Column 3 counts involvement in an activity only if it generates at least $60 over the previous 12 months for the household.

Table 2: Determinants of Entrepreneurship Choice

A: Nicaragua: Weighted probit[a]

	(1)		(2)		(3)		(4)	
	dP	t[b]	dP	t	dP	t	dP	t
Household characteristics								
Number of male adults			3.08	2.23	3.00	2.24	3.42	2.55
Number of female adults			2.55	2.01	2.82	2.30	2.71	2.20
Average age			4.07	2.95	3.89	2.88	3.83	2.89
Human capital index			0.83	0.58	0.84	0.59	1.06	0.75
Head female			6.14	2.06	5.31	1.81	5.74	1.98
Head parents entrepreneur			2.27	0.86	2.06	0.80	2.19	0.86
ln(Other income)			-2.66	2.08	-2.82	2.21	-2.42	1.91
ln(Remittances)			-7.53	5.83	-7.26	5.71	-7.27	5.71
ln(Assets)			5.14	3.59	4.99	3.56	4.71	3.40
Benchmarks and components								
Connectivity	-0.80	0.38	-0.42	0.19	-0.53	0.23		
Proximity to post office							2.78	2.07
Infrastructure services	5.45	2.32	5.78	2.44	2.78	1.10		
Access to water							3.33	2.50
Business services	-1.91	1.25	-2.28	1.48	-1.80	1.16		
Governance	-0.74	0.59	-0.45	0.35	0.09	0.07		
Human capital	7.91	4.90	6.29	3.79	4.89	2.45	4.49	2.70
Finance services	-0.37	0.25	-1.00	0.69	-1.29	0.85		
Community variables								
ln(Community size)					0.78	0.36		
ln(Income per capita)					2.76	1.36		
Agricultural seasonality					1.94	1.40		
Enterprise openness					3.67	2.48	5.30	3.91
ln(Male wage)					-0.48	0.35		
Regression statistics								
Log-Likelihood	-580.19		-544.10		-535.83		-535.90	
Number of observations	1163		1163		1163		1163	

B: Sri Lanka: Weighted random effect probit[a]

	(1)		(2)		(3)		(4)	
	dP	t	dP	t	dP	t	dP	t
Household characteristics								
Number of male adults			2.73	1.26	2.92	1.38	3.25	1.54
Number of female adults			-2.17	1.00	-2.03	0.94	-1.88	0.88
Average age			0.91	0.33	0.84	0.32	0.58	0.23
Human capital index			0.93	0.34	1.04	0.38	1.57	0.60
Head female			-8.61	1.03	-8.83	1.08	-10.40	1.42
Not Sinhalese			-4.99	0.51	-4.87	0.51	-14.24	1.80
Head parents entrepreneur			1.56	0.31	2.11	0.42	2.09	0.43
ln(Assets)			9.74	3.13	9.82	3.18	10.55	3.59
Benchmarks and components								
Connectivity	5.70	1.29	5.34	1.28	4.29	1.04		
Proximity to post office							4.84	1.84

Infrastructure services	-2.19	0.64	-2.08	0.58	-4.14	1.13		
Sewage system							-7.33	2.54
Business services	3.49	1.13	3.66	1.27	3.36	1.08		
Engineering services							-4.49	1.69
Information technology services							5.93	1.58
Governance	2.89	1.07	3.91	1.43	3.75	1.40		
Human capital	4.83	1.35	2.41	0.60	3.36	0.81		
Finance services	-2.95	0.67	-3.68	0.83	-4.26	1.02		
Community variables								
ln(Community size)					4.44	1.25	3.98	1.35
ln(Income per capita)					0.90	0.31		
Agricultural seasonality					-0.96	0.27		
Enterprise openness					-0.72	0.20		
ln(Male wage rate)					2.59	0.78	4.36	1.67
Regression statistics								
Standard error of random eff.	10.38	3.80	9.92	3.65	9.80	3.49	6.99	2.81
Log-Likelihood	-474.70		-454.90		-451.39		-438.06	
Number of observations	849		849		849		849	

Notes: [a] Columns headed with "dP" report the percentage point increase in the probability that an average household operates an enterprise in response to a one-standard-deviation increase in the explanatory variable. For the variables "Head female", "Head parents entrepreneur", and "Not Sinhalese", which are dummy variables, the change in the explanatory variable is one unit.

[b] Significance levels implied by the t-statistics are: 1% if $t \geq 2.58$, 5% if $1.96 \leq t < 2.58$, 10% if $1.645 \leq t < 1.96$.

Source: RICS data

Table 3: Contributions to explanation of entrepreneurship choice

Explanatory variables	*Nicaragua*		*Sri Lanka*	
	Log-likelihood value	Increment as % of base	Log-likelihood value	Increment as % of base
None (base)	–611.797		–489.841	
A. Household characteristics only	–562.312	8.09%	–467.398	4.58%
+ Community random effect	–562.312	0.00%	–462.053	1.09%
+ Benchmark indicators	–544.097	2.98%	–454.898	1.46%
+ Community characteristics	–535.835	1.35%	–451.391	0.72%
Total		12.42%		7.85%
B. Household characteristics only	–562.312		–467.398	
+ Community random effect and best components	–535.900	4.32%	–438.063	5.88%
C. Benchmark indicators and community random effect only	–580.186	5.17%	–474.698	3.09%
D. Community characteristics and community random effect only	–579.148	5.34%	–479.544	2.10%

Source: Author's calculation from RIC data

Table 4: Determinants of Household Activity Choices

A: Nicaragua: Weighted probit and weighted random effect probit[a]

	Non-farm household enterprise		Wage employment		Non-farm economic activity		Farming	
	dP	t[b]	dP	t	dP	t	dP	t
Household characteristics								
Number of male adults	3.42	2.55	7.91	3.03	3.41	1.80	4.01	1.40
Number of female adults	2.71	2.20	1.47	0.69	1.03	0.71	0.16	0.05
Average age	3.83	2.89	-5.46	2.47	-1.61	1.17	10.15	3.64
Human capital index	1.06	0.75	9.11	4.43	7.07	4.26	-10.62	3.36
Head female	5.74	1.98	2.60	0.52	0.53	0.15	-6.16	0.90
Head parents entrepreneur	2.19	0.86	1.69	0.42	2.03	0.67	-8.89	1.37
ln(Other income)	-2.42	1.91	-4.86	2.32	-4.80	3.99	-4.59	1.61
ln(Remittances)	-7.27	5.71	-6.42	3.13	-7.89	5.50	-4.32	1.36
ln(Assets)	4.71	3.40	-18.77	6.87	-9.59	4.92	19.33	5.47
Benchmarks and components								
Proximity to post office[c]	2.78	2.07						
Percent of households with electricity[d]			-8.18	2.15	-4.84	1.42		
Infrastructure services to water[d]	3.33	2.50						
Sewage system[d]							-8.97	1.47
Dealing with government agencies[f]			-7.19	2.06	-4.09	1.42		
Rule of law[f]							-7.13	1.31
Human capital	4.49	2.70	-6.72	1.22				
Community variables								
ln(Income per capita)			11.01	1.99	5.47	1.31	-15.14	2.56
Agricultural seasonality			-7.31	1.72			8.37	1.62
Enterprise openness	5.30	3.91			3.87	1.15		
Regression statistics								
Standard error of random effect	0.00		15.46	5.41	12.00	4.41	19.07	5.12
Log-Likelihood	-535.90		-544.73		-330.56		-498.38	
Number of observations	1163		1163		1163		1163	

B: Sri Lanka: Weighted random effect probit[a]

	Non-farm household enterprise		Wage employment		Non-farm economic activity		Farming	
	dP	t	dP	t	dP	t	dP	t
Household characteristics								
Number of male adults	3.25	1.54	5.32	2.27	0.64	0.75	12.26	4.08
Number of female adults	-1.88	0.88	10.10	4.72	2.66	3.49	-4.40	1.45
Average age	0.58	0.23	-5.41	2.45	-2.44	3.05	4.84	1.69
Human capital index	1.57	0.60	7.68	3.20	3.29	3.57	-3.09	0.91
Head female	-10.40	1.42	12.17	2.34	-1.68	0.95	-7.50	0.87
Not Sinhalese	-14.24	1.80	4.33	0.49	-5.85	2.23	17.76	1.39
Head's parents entrepreneur	2.09	0.43	-0.21	0.04	2.19	1.21	-1.89	0.27
ln(Assets)	10.55	3.59	-6.04	2.13	-0.42	0.47	20.64	4.87
Benchmarks and components								

Inverse cost of transport to major city[c]			-6.55	1.65				
Inverse cost of transport to main market[c]			5.73	1.63				
Proximity to post office[c]	4.84	1.84						
Infrastructure services			5.61	1.78	2.28	1.76		
Percent of households with electricity[d]							-8.54	1.32
Percent of households with fixed phone line[d]							-13.38	1.93
Sewage system[d]	-7.33	2.54						
Engineering services[e]	-4.49	1.69						
Management consulting services[e]							-15.62	2.17
Marketing services[e]			4.94	2.18			18.92	2.75
Information technology services[e]	5.93	1.58						
Conflict resolution, contract enforcement[f]					1.67	1.37		
Financial penetration[g]			-6.38	2.38			9.32	1.56
Community variables								
ln(Community size)	3.98	1.35	4.90	1.66	3.35	2.34	-11.52	1.88
Agric. Seasonality							9.20	1.36
ln(Male wage rate)	4.36	1.67						
Regression statistics								
Standard error of random effect	6.99	2.81	8.79	3.85	5.38	2.86	24.81	5.18
Log-Likelihood	-438.06		-390.69		-188.94		-401.02	
Number of observations	849		849		849		849	

Notes: [a] Columns headed with "dP" report the percentage point increase in the probability that an average household operates an enterprise in response to a one-standard-deviation increase in the explanatory variable. For the variables "Head female", "Head parents entrepreneur", and "Not Sinhalese", which are dummy variables, the change in the explanatory variable is one unit.

[b] Significance levels implied by the t-statistics are: 1% if $t \geq 2.58$, 5% if $1.96 \leq t < 2.58$, 10% if $1.645 \leq t < 1.96$.

[c] A component of the connectivity index.

[d] A component of the infrastructure services index.

[e] A component of the business services index.

[f] A component of the governance index.

[g] A component of the finance services index.

Source: RICS data

Table 5: Determinants of Enterprise Start-Up[a]

A: Nicaragua

	(1)		(2)		(3)		(4)	
	dP	t	dP	t	dP	t	dP	t
Household characteristics								
Number of male adults			-0.04	0.13	-0.10	0.37	-0.08	0.29
Number of female adults			0.08	0.21	0.12	0.31	0.14	0.37
Average age			-0.39	1.05	-0.50	1.31	-0.41	1.13
Human capital index			0.14	0.42	0.23	0.74	0.21	0.75
Head female			1.15	1.40	1.09	1.32	1.05	1.33
Head's parents entrepreneur			-0.10	0.14	-0.18	0.25	-0.06	0.09
ln(Other income)			-0.61	1.81	-0.68	2.03	-0.67	1.96
ln(Remittance income)			-1.08	2.85	-0.99	2.74	-0.94	2.65
ln(Assets)			0.31	1.05	0.33	1.09	0.36	1.18
Benchmarks and components								
Connectivity	-0.96	1.60	-0.89	1.55				
Inverse cost of transport to major city							-0.57	1.62
Infrastructure services	0.72	1.29	0.79	1.45				
Percent of households with fixed phone line							-0.52	1.20
Business services	0.26	0.51	0.13	0.31				
Governance	-0.35	1.04	-0.35	1.14				
Human capital	0.54	1.41	0.38	0.99				
Finance	0.26	0.65	0.27	0.74				
Community variables								
ln(Community size)					0.45	1.16	0.72	1.83
ln(Income per capita)					-0.17	0.39		
Agricultural seasonality					0.50	1.52	0.40	1.24
Enterprise openness					0.64	1.96	0.63	2.11
ln(Male wage rate)					0.44	1.51		
Regression statistics								
lnL	-156.51		-152.46		-151.34		-150.91	
n	1170		1170		1170		1170	

B: Sri Lanka

	(1)		(2)		(3)		(4)	
	dP	t	dP	t	dP	t	dP	t
Household characteristics								
Number of male adults			-0.11	0.66	-0.19	1.13	-0.16	1.06
Number of female adults			0.20	1.45	0.15	1.19	0.17	1.38
Average age			-0.27	1.60	-0.26	1.48	-0.25	1.52
Human capital index			-0.17	0.93	-0.11	0.71	-0.21	1.18
Head female			0.28	0.76	0.33	0.88	0.26	0.77
Not Sinhalese			-1.05	1.68	-0.77	1.37	-1.04	1.89
Head parents entrepreneur			0.24	0.66	0.51	1.41	0.37	1.06
ln(Assets)			0.43	2.11	0.46	2.28	0.43	2.38
Benchmarks and components								
Connectivity	-0.02	0.10	0.03	0.13				
Inverse cost of transport to major city							0.33	1.81

Inverse cost of transport to main market							-0.26	1.65
Infrastructure services to infrastructure	-0.37	1.46	-0.47	1.96			-0.40	2.27
Business services	0.19	0.83	0.21	1.06				
Governance	-0.17	1.20	-0.17	1.28				
Human capital	0.38	2.02	0.44	1.93			0.28	1.47
Finance	0.10	0.53	0.08	0.42				
Community variables								
ln(Community size)					-0.16	0.96		
ln(Income per capita)					0.00	0.02		
Agric. Seasonality					0.23	1.26		
Enterprise openness					-0.29	1.71	-0.30	1.88
ln(Male wage rate)					0.36	2.05	0.24	1.38
Regression statistics								
lnL	-50.54		-49.29		-49.00		-48.48	
n	790		790		790		790	

Notes: [a] Estimates of a weighted probit model. Columns headed with "dP" report the percentage point increase in the probability that an average household operates an enterprise in response to a one-standard-deviation increase in the explanatory variable. For the variables "Head female", "Head parents entrepreneur", and "Not Sinhalese", which are dummy variables, the change in the explanatory variable is one unit.

[b] Significance levels implied by the t-statistics are: 1% if $t \geq 2.58$, 5% if $1.96 \leq t < 2.58$, 10% if $1.645 \leq t < 1.96$.

Source : RICS data