

IZA DP No. 8524

**Labor Productivity in Rural African Enterprises:  
Empirical Evidence from the LSMS-ISA**

Paula Nagler  
Wim Naudé

October 2014

# Labor Productivity in Rural African Enterprises: Empirical Evidence from the LSMS-ISA

**Paula Nagler**

*Msm, UNU-MERIT/MGSoG, Maastricht University*

**Wim Naudé**

*Msm, UNU-MERIT/MGSoG, Maastricht University  
and IZA*

Discussion Paper No. 8524  
October 2014

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## **ABSTRACT**

### **Labor Productivity in Rural African Enterprises: Empirical Evidence from the LSMS-ISA**

Although non-farm enterprises are ubiquitous in rural Sub-Saharan Africa, little is yet known about their productivity. In this paper we contribute to filling this gap by providing estimates of labor productivity in enterprises for Ethiopia, Malawi, Nigeria, and Uganda. Using the World Bank's LSMS-ISA database, we find that rural enterprises are on average less productive than those in urban areas, and that female-owned enterprises are less productive than male-owned enterprises. By estimating Heckman selection and panel data models, we find that education and access to credit are associated with higher labor productivity, while households that experience shocks operate less productive enterprises. Furthermore we provide evidence that enterprises that operate throughout the year are more productive. We conclude that gender, education, shocks, access to finance, and location matter for labor productivity in rural Africa, and that policy decisions tackling the shortcomings could significantly contribute to a better business environment and increased labor productivity.

JEL Classification: J43, L26, M13, O13, O55, Q12

Keywords: entrepreneurship, informal sector, labor productivity, rural development, Sub-Saharan Africa

Corresponding author:

Wim Naudé  
Maastricht School of Management  
PO Box 1203  
6201 BE Maastricht  
The Netherlands  
E-mail: [w.naude@maastrichtuniversity.nl](mailto:w.naude@maastrichtuniversity.nl)

# 1 Introduction

While non-farm enterprises are ubiquitous in rural Africa and play an important role in the non-farm economy, little is yet known about their productivity.<sup>1</sup> Almost one in two households operates such an enterprise and provides a large variety of products and services to the rural population (Nagler and Naudé, 2014). The relative contribution of the non-farm economy to household income is significant and has increased over time (Haggblade et al., 2010; Rijkers and Costa, 2012). These enterprises are further estimated to currently employ 15 percent of Africa’s labor force (Fox et al., 2013) and expected to create millions of jobs in rural Africa over the next decade (Fox and Pimhidzai, 2013). They are also expected to contribute to structural transformation (Fox and Sohnesen, 2013), although the non-farm sector largely consists of small, informal household businesses. Nevertheless, the literature so far lacks empirical evidence on the productivity of these enterprises. In contrast, a number of studies have already investigated the productivity of urban enterprises in Africa (e.g. Alene, 2010; Block, 2010), including the impact of non-farm enterprises on agricultural productivity (e.g. Oseni and Winters, 2009). This knowledge gap on rural enterprise productivity presents a serious lacuna, as more productive enterprises have a higher likelihood to survive and to contribute towards rural employment creation (Wennberg and Lindqvist, 2010; Owoo and Naudé, 2014).

In this paper we contribute to the literature by exploring the productivity of non-farm enterprises in four African countries using the World Bank’s recent LSMS-ISA data set. The database covers six countries in Sub-Saharan Africa over the period 2005 to 2013. We focus on Ethiopia, Malawi, Nigeria, and Uganda due to data availability. To the best of our knowledge we are the first to provide empirical evidence of labor productivity in African enterprises using this database. Based on the literature and our companion paper (Nagler and Naudé, 2014) we expect that factors pushing households into entrepreneurship have a negative impact on labor productivity, and that factors pulling households into entrepreneurship to seize opportunities have a positive impact on labor productivity. We use Heckman selection and panel data estimators, and take labor productivity, a partial productivity measure, as a proxy for enterprise performance.

The results confirm our expectations that push factors decrease and pull factors increase labor productivity, and that location matters. Rural enterprises are less productive than their urban counterparts, and enterprises that are located closer to a population center are more productive than enterprises located further away. We further show that female-owned enterprises are on average less productive than male-owned enterprises, and that age increases labor productivity, a proxy for experience. External shocks have a negative effect on labor productivity, indicating that households pushed into entrepreneurship operate less productive businesses. Education (proxied by the ability to read & write) and months of enterprise operation have a positive effect on labor productivity, indicating that pull factors increase enterprise performance.

---

<sup>1</sup> We use the term “rural non-farm enterprises” to refer to generally small, informal household enterprises in the rural non-farm economy. These enterprises include all non-agricultural activities in rural areas including on-farm (but non-agricultural) activities such as agribusiness, as well as services, trade and retail, tourism, rural industrialization, construction, and mining (Nagler and Naudé, 2014).

The paper is structured as follows. In section 2 we review the theoretical and empirical literature of enterprise productivity. In section 3 we describe the LSMS-ISA database and provide descriptive statistics on the dispersal of labor productivity. In section 4 we present the methodology used, and in section 5 we discuss our empirical results. The final section summarizes the findings and discusses policy implications.

## 2 Literature Review

The literature review is organized in two broad sections. In section 2.1 we survey the theoretical literature to identify determinants of enterprise productivity, and in section 2.2 we report relevant empirical literature focusing on Sub-Saharan Africa.

### 2.1 Determinants of Enterprise Productivity

An extensive review of the literature on determinants of enterprise productivity is provided by Syverson (2011). The large majority of studies have analyzed enterprise performance in developed economies and identified education, work experience and skills levels of managers and workers, and market structure as the main determinants of enterprise productivity (Moretti, 2004; van Biesebroeck, 2005; Bloom and Reenen, 2010). The literature also finds that female-headed enterprises are less productive than male-headed enterprises (Amin, 2011; Kinda et al., 2011; Saliola and Seker, 2011; Rijkers and Costa, 2012). Furthermore enterprises clustered together are more productive due to localization and urbanization economies, such as knowledge and technology spill-overs (Martin et al., 2011; Bloom et al., 2013), horizontal linkages (Nichter and Goldmark, 2009), and more competition (Foster et al., 2008; Ali and Peerlings, 2011). As a result the geographical location of an enterprise, and in particular the proximity to other productive enterprises have an impact on its productivity.

While we expect these determinants to be potentially relevant in the case of rural enterprises in Africa, a major difference to consider is the lack of a watertight demarcation between the family and the enterprise, as assumed in micro-economic theory and applied in empirical work of developed countries. In developed economies, most determinants of entrepreneurship and firm productivity relate to the individual entrepreneur and to firm characteristics. In contrast, rural enterprises tend to be household enterprises, and hence entrepreneurial decisions largely taken at the household level. Most households are furthermore involved in agriculture, meaning that the choice to operate a non-farm enterprise is not only determined at the household level, but equally intertwined with farming decisions.<sup>2</sup>

---

<sup>2</sup> The general entrepreneurship literature has neglected the fact of household level decision-making in case of family firms, focusing most of the time on the individual entrepreneur as the agent. As Alsos et al. (2013, p.10) recognize, household level decisions are “influential on how businesses are started and managed, but rarely garner attention within the entrepreneurship research literature”.

In this context, the implications from agricultural household models are taken into account. Following [Singh et al. \(1986\)](#), rural households in Africa are treated as both consumers and producers, assuming complete markets and information ([Taylor and Adelman, 2003](#)). However, rural markets are far from perfect ([Janvry and Sadoulet, 2006](#)). In addition to missing, incomplete or imperfect markets, households face different types of external shocks, such as climatic or economic shocks, diseases, and conflicts ([Sabates-Wheeler et al., 2013](#)). Once market failures and risks are taken into consideration, the agricultural households' desire to smooth consumption, diversify income, or cope with shocks in the context of incomplete insurance and credit markets, can drive them to create a non-farm enterprise as a risk-coping and income diversification strategy ([Janvry and Sadoulet, 2006](#); [Dercon, 2009](#)). As a consequence, households may choose to operate a business that can easily be started and exited, operated occasionally, and that provides mainly employment for household members. These kind of businesses are then likely to be enterprises of low productivity and intermittent operation in the informal sector (see also [Rosenzweig and Binswanger, 1993](#); [Janvry and Sadoulet, 2006](#); [Loening et al., 2008](#)).

As a result, we expect to find a link between a household's motivation to operate a non-farm enterprise and its subsequent productivity. We assume that enterprises that are operated by necessity are more likely to be less productive than enterprises operated as a consequence to perceiving an opportunity. The latter may not only report higher productivity levels due to attaining better capacity utilization by operating all year long, but also because households may be able to access credit or pursue training and education, all factors associated with higher productivity in the literature. An additional factor that may pull rural households into operating an enterprise is location. On the one hand location is affected by weather conditions such as climate and rainfall that influence agricultural productivity, on the other hand by economic geography such as distance from an urban center and market access, reflecting demand and competition.

From these theoretical considerations we expect that the determinants of labor productivity of rural African enterprises are related to the household's motivation to operate an enterprise, reflected by agricultural risk and productivity, economic geography, external shocks, and facilitating characteristics on both the individual and household level such as gender, age, education, access to credit, and number of adults within the household.

## **2.2 Enterprise Productivity in Sub-Saharan Africa**

While a solid body of literature has dealt with the above dimensions of enterprise productivity in developed countries, only a few studies have analyzed enterprise productivity in Sub-Saharan Africa. Most of those that do, focus either on formal enterprises or manufacturing enterprises and tend to be overwhelmingly urban-based. Empirical evidence on the performance of non-farm enterprises in rural Africa is however scarce. This lack of knowledge is a serious lacuna, since policies to support rural development have the potential to substantially enhance their effectiveness if they also succeed in raising the productivity of these enterprises.

For African countries most work has been conducted on the impact of the business environment on enterprise productivity in the formal sector. Various studies, using World Bank data on these formal enterprises, have found that a poor business environment reduces enterprise productivity and growth. Studies include [Dollar et al. \(2005\)](#), [Arnold et al. \(2006\)](#), [Eifert et al. \(2008\)](#), and [Dethier et al. \(2010\)](#). Less work has been done on clustering and spatial spill-overs. The only study for rural Africa so far conducted, uses geo-referenced household data from Ethiopia and Nigeria and finds evidence of significant spatial auto-correlation from the productivity levels of enterprises in one location, the enumerator area, on that of closely located enterprises ([Owoo and Naudé, 2014](#)).

The remaining studies have focused on manufacturing firms in Africa. [Rijkers et al. \(2010\)](#) analyze the productivity of manufacturing enterprises in rural Ethiopia. The authors find that rural enterprises are less productive than urban ones, and report an output per labor ratio for remote rural enterprises of 0.43, while it is 0.95 for enterprises in rural towns, and 2.30 for enterprises in urban areas ([Rijkers et al., 2010](#), p. 1282). Furthermore they point out that productivity levels are more dispersed in the case of rural enterprises, and that female-headed enterprises are less productive than male-headed enterprises. [Söderbom and Teal \(2004\)](#) and [Söderbom et al. \(2006\)](#) find that more productive firms tend to survive longer in Africa. This happens however only in the cases when these enterprises have already attained a certain firm size. [Frazer \(2005\)](#) confirms this finding, using Ghanaian enterprise-level data, that more productive enterprises are more likely to survive compared to less productive ones.

### 3 Database and Descriptive Statistics

In this section we first describe the database, the World Bank’s Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA), and provide an overview of the shortcomings. Second, we outline the methodology used in the paper to estimate determinants of labor productivity.

#### 3.1 The Database

The LSMS-ISA database is the result of nationally representative, cross-sectional and longitudinal surveys conducted by the World Bank in collaboration with national statistical offices in various countries in Sub-Saharan Africa.<sup>3</sup> The surveys cover six countries at the time of writing: Ethiopia, Malawi, Niger, Nigeria, Tanzania, and Uganda. Cross-sectional data is currently available for all countries, and panel data for Nigeria, Tanzania, and Uganda.

The surveys consist of three parts: a community questionnaire, an agricultural questionnaire and a household questionnaire. The community questionnaire collects

---

<sup>3</sup> Additional information on the LSMS-ISA database: [www.worldbank.org/lsms-isa](http://www.worldbank.org/lsms-isa)

community-level information and covers access to public services and infrastructure, social networks, governance, and retail prices. The agricultural questionnaire collects information on crop production, storage and sales, land holdings, farming practices, input use and technology adaption, access to and use of services, infrastructure and natural resources, livestock, and fishery. Finally, the household questionnaire captures household demographics, migration, education, health and nutrition, food consumption and expenditure, non-food expenditure, employment, non-farm enterprises and further income sources, dwelling conditions, durable assets, and participation in projects and programs. All data is furthermore geo-referenced.

Despite its strengths and usefulness, the data set has some weaknesses. A comprehensive analysis of enterprise survival and failure is constrained due to a lack of data on failed enterprises. Furthermore, enterprises are not explicitly tracked over the survey rounds in Tanzania and Uganda. Therefore we cannot follow enterprises to study their individual performance over time, and empirical estimates using duration analysis are not possible. More general, but less precise estimations, are still possible. Following our empirical analysis in the next section, we suggest that in future waves and countries, where the LSMS-ISA surveys is still planned to be conducted, a number of modifications and additions should be made to the questionnaire. Although a full elaboration of these fall outside the scope of the present paper, we can briefly mention some to illustrate the type of current shortcomings we face. First, apart from clearly tracking individual enterprises over the survey rounds, discontinued enterprises should equally be registered and information collected. We suggest to add specific questions to know more about these discontinued enterprises, e.g. when exactly the enterprise was discontinued, and the cause for terminating enterprise operations. Second, we recommend to add modules to the questionnaires that collect information to estimate production functions and technical efficiency, including price information, information about the final product(s) and their prices, wages paid per worker, input costs (intermediate inputs), overhead (transport, security), as well as fixed and working assets employed in the enterprise to know more about the capital intensity.

## 3.2 Productivity Dispersal of Non-Farm Enterprises

In this section we describe the patterns of labor productivity dispersal by location, gender, education, and the experience of a shock. Enterprise productivity refers to how efficiently enterprises transform inputs into outputs. The most frequently used measures are total factor productivity and partial measures such as output per labor unit. In the present case, a lack of data precludes calculation of total factor productivity. Hence we can only calculate labor productivity, a partial productivity measure.

Formally we calculate,

$$\text{labor productivity} = \frac{\text{average monthly sales}}{\text{number of workers}}$$

As average monthly sales is not available for all countries, we take total sales during the



last month of operation in Malawi and Nigeria, and average gross revenues in Uganda. Once labor productivity is calculated, we take the log of labor productivity for our estimates. We use the data survey rounds from 2011/12 for Ethiopia, and from 2010/11 for Malawi, Nigeria, and Uganda.

We acknowledge that using this productivity measure is an approximation, and likely to be subject to measurement errors in the present context. First, we do not take into account output or profit due to a lack of data. Second, we do not have information about the time-use of workers, and do not know how many hours of work are allocated by each individual to the enterprise. And third, we assume that the self-reported variable itself is imprecise, as most enterprises are informal and do not always record sales or revenues, but provide rough estimates of the business volume. Therefore results have to be interpreted with caution. Due to data availability, this productivity measure is however the best approximation that is currently at hand for this study.

We depict the dispersal by location in Figure 1. The Figure shows that differences in labor productivity confirm our expectations: urban enterprises are more productive than rural enterprises in the four countries we study, with a more pronounced productivity gap in Malawi and Uganda.<sup>4</sup>

However, this simple dichotomy between rural and urban areas may be of limited usefulness in Africa. There is a large variation in Africa's economic geography between deep rural, small towns, and major urban areas. The potential importance of secondary towns and rural agglomerations has generally been underestimated. As [Christiaensen et al. \(2013\)](#) argue, productivity might be higher in metropolitan areas, but not all inhabitants from rural areas are or will be able to access opportunities in these areas. Using data from Tanzania they find that only one in seven people who escaped from poverty did so through migrating to a large city, but that one in two did so by moving to a secondary town.

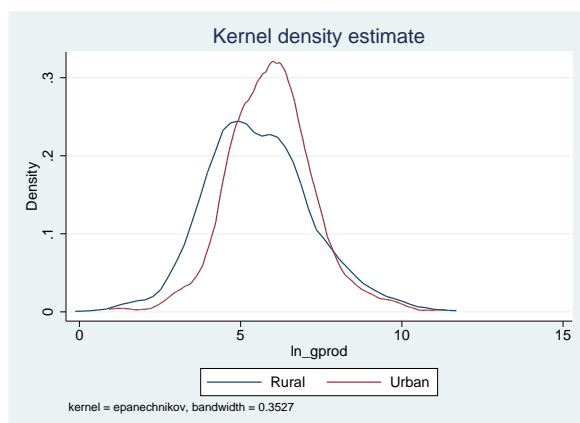
Therefore we are interested in providing a finer analysis of how spatial location affects productivity of non-farm enterprises. We use data from Uganda and analyze it with regard to the country's division into four main regions and the capital city: Central, Eastern, Northern, Western, and the capital Kampala. These four regions plus the capital city are distinct in terms of population density, business environment, and history. Over the past ten years economic growth has been largely concentrated in the Central region that includes the capital city, while violent conflict affected the Northern region between 1987 and 2006 ([Blattmann et al., 2014](#)). The Western region also has not escaped occasional (ethnic) conflicts ([Espeland, 2007](#)). This allows us to obtain some indication of how the local business environment may affect labor productivity, as regions that experienced a conflict are more likely to show a poorer business environment ([Brueck et al., 2013](#)). We also expect that labor productivity is higher in the capital city, and the more densely populated and more peaceful regions. Figure 2a confirms our expectations. Kampala has the highest labor productivity, with the Central region following in terms of productivity level. The Northern region, with the lowest population density and a history of conflict,

---

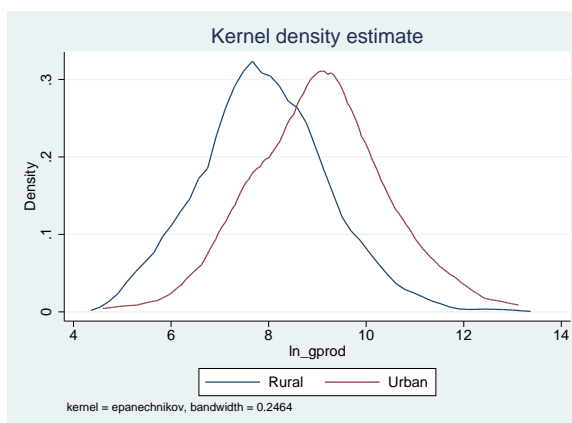
<sup>4</sup> The data does not cover urban areas in Ethiopia. Instead the differentiation is made between rural areas and small towns. We assume that the small difference in productivity dispersal in Ethiopia stems from this data limitation.

**Figure 1: Productivity Dispersal by Location**

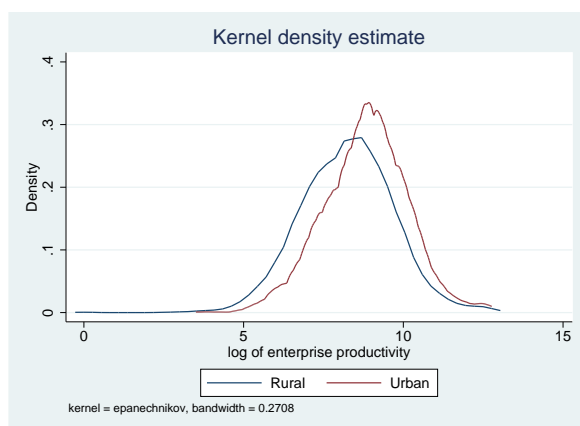
(a) Ethiopia



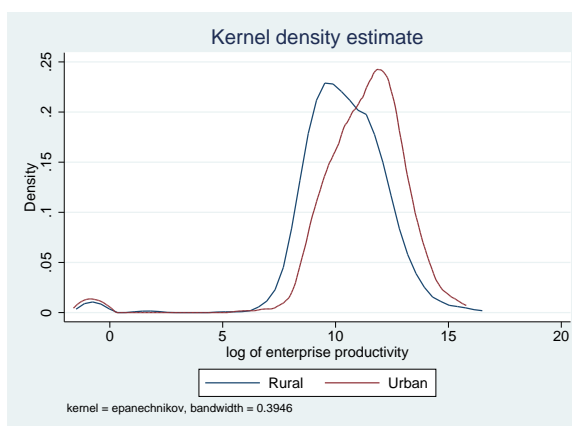
(b) Malawi



(c) Nigeria



(d) Uganda



Source: Authors' own calculations based on the LSMS-ISA database.

is home to the enterprises with the lowest productivity level.

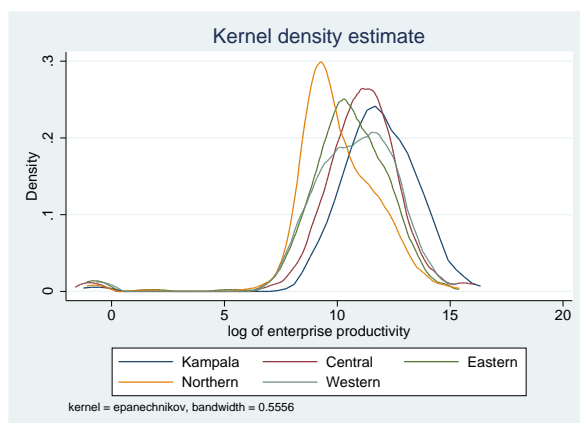
We also consider the impact of distance to the closest town on labor productivity in non-farm enterprises. We use again data from Uganda and take different distances of the enterprise location to the closest population center.<sup>5</sup> We compare the productivity density of the enterprises by various distance categories from these population centers. We expect, following [Christiaensen et al. \(2013\)](#), that proximity to a population center increases labor productivity and that it gradually diminishes over distance, reflecting that spatial spillovers are limited. Figure 2b confirms our expectations. Enterprises of households located up to 10 km from a population center are the most productive ones, followed by enterprises in households residing up to 25 km and 50 km away, respectively. If households are located more than 50 km away, the results show a significantly lower labor productivity. These findings support the idea that rural secondary towns and cities, providing links between rural areas and major cities, may play an important role in the structural transformation and poverty reduction of agrarian societies like Uganda.

We further depict the productivity dispersal by gender, education (proxied by the ability

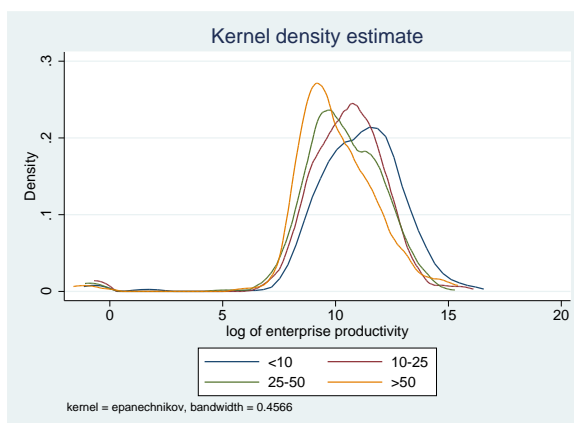
<sup>5</sup> Population center is defined as a town or city with 20,000 or more inhabitants.

**Figure 2:** Productivity Dispersal by Region and Distance

(a) Uganda - by Region



(b) Uganda - by Distance (in km)



Source: Authors' calculations based on LSMS-ISA data.

to read & write), and the experience of shocks in Figures 3, 4 and 5.<sup>6</sup> Enterprises with a male owner are on average more productive than enterprises with a female owner in all four countries. A possible explanation could lie in the fact that businesses are time-consuming to operate, and women tend to be more time-constrained due to household duties (see also Palacios-López and López, 2014). The productivity dispersal by the ability to read & write shows that literate enterprise owners operate more productive enterprises. However, the ability to read & write is an imprecise approximation to the educational level of the entrepreneur, capturing individuals with primary education, but also individuals who have more comprehensive education, for example secondary schooling. Finally, we report the productivity dispersal by the experience of a shock during the last 12 months preceding the survey. As expected, we find that external shocks to the household show a negative association with labor productivity. Labor productivity is higher in enterprises where the household did not experience a shock. While the difference is minor in Ethiopia and Nigeria, it is more pronounced in Malawi and Uganda.

## 4 Methodology

In this section we present the methodology used to estimate labor productivity in non-farm enterprises. We apply two econometric models: a Heckman selection model that accounts for selection effects, and a panel data model, where we make use of the panel data that is available for Nigeria and Uganda.

### 4.1 Heckman Selection Model

We use a Heckman selection model to estimate the determinants of labor productivity, since the motivation for operating a business can determine its subsequent productivity

<sup>6</sup> Figures are reported in Appendix B.

and the results therefore subject to selection effects if not accounted for. For instance, if households are pushed into entrepreneurship as a risk-management tool due to shocks and uncertainty in farming, other factors might determine labor productivity compared to a situation, where households perceive an attractive business opportunity. This model takes into account the difference between the probability that a particular household operates a non-farm enterprise, called the selection stage, and the level of labor productivity for this subsequent enterprise, called the outcome stage. To estimate the model we use a cross-sectional sample, where we include data from the four countries, taking the same survey rounds as indicated in section 3.2.

Formally we estimate,

$$z_i^* = w_i\gamma + u_i \tag{1}$$

representing the selection stage of the model, where  $z_i^*$  is a latent variable corresponding to the “outcome” productivity level, which will only be observed once a household operates a non-farm enterprise, thus  $z_i = 1$  if  $z_i^* > 0$  and  $z_i = 0$  if  $z_i^* \leq 0$ .  $w_i$  is a vector containing the possible determinants of enterprise operation.

Once  $z_i$  is known, the outcome stage, the dependent variable is the log of labor productivity and can be modeled as,

$$y_i^* = x_i\beta + \varepsilon_i \tag{2}$$

with  $y_i = y_i^*$  if  $z_i = 1$  and  $y_i$  not observed if  $y_i = 0$ .

$x_i$  is a vector including the individual, household, location and enterprise characteristics. In the selection stage we take the individual characteristics of the household head, and include the variables gender, age, and education. In the outcome stage we take the individual characteristics of the enterprise owner, and include the same variables as in the selection stage. The households characteristics, general access to credit and experience of shocks, remain equal in both stages, as well as the location characteristics, distance to the next population center and agro-ecological zones. Enterprise characteristics are only included in the outcome stage, where we add firm size to the regressions.<sup>7</sup>

## 4.2 Panel Data Analysis

We make use of the panel properties for Nigeria and Uganda to estimate panel regressions. For the panel data analysis we have data for Nigeria for four survey rounds: post-planting and post-harvest data for the years 2010/11 and 2012/13. Due to a limited questionnaire for the post-harvest interviews, we use the data of the post-planting rounds only.<sup>8</sup> In the case of Uganda, we have data for three survey rounds: 2009/10, 2010/11, and 2011/12.

---

<sup>7</sup> See Appendix C.1 for summary statistics and a detailed definition of the variables used in the Heckman selection model.

<sup>8</sup> The shock variable is taken from the post-harvest questionnaires due to data availability.

Formally we estimate,

$$Y_{eht} = \mathbf{x}'_{eht}\boldsymbol{\beta}^1 + \mathbf{w}'_{ht}\boldsymbol{\gamma}^1 + \mathbf{z}'_{eht}\boldsymbol{\delta}^1 + \lambda_t^1 + \mu_h^1 + u_{eht}^1 \quad (3)$$

$$\bar{Y}_{eht} = \bar{\mathbf{x}}'_{eht}\boldsymbol{\beta}^2 + \bar{\mathbf{w}}'_{ht}\boldsymbol{\gamma}^2 + \bar{\mathbf{z}}'_{eht}\boldsymbol{\delta}^2 + \lambda_t^2 + \mu_h^2 + u_{eht}^2 \quad (4)$$

$$\bar{\bar{Y}}_{eht} = \bar{\bar{\mathbf{x}}}'_{eht}\boldsymbol{\beta}^3 + \bar{\bar{\mathbf{w}}}'_{ht}\boldsymbol{\gamma}^3 + \bar{\bar{\mathbf{z}}}'_{eht}\boldsymbol{\delta}^3 + \lambda_t^3 + \mu_h^3 + u_{eht}^3 \quad (5)$$

where  $\{Y_{eht}, \bar{Y}_{eht}, \bar{\bar{Y}}_{eht}\}$  is labor productivity at the enterprise level. The set  $\{\mathbf{x}'_{eht}, \bar{\mathbf{x}}'_{eht}, \bar{\bar{\mathbf{x}}}'_{eht}\}$  contains vectors of individual characteristics of the enterprise owner at the enterprise level. The set  $\{\mathbf{w}'_{ht}, \bar{\mathbf{w}}'_{ht}, \bar{\bar{\mathbf{w}}}'_{ht}\}$  contains vectors of household characteristics observed at the household level. Finally,  $\{\mathbf{z}'_{eht}, \bar{\mathbf{z}}'_{eht}, \bar{\bar{\mathbf{z}}}'_{eht}\}$  is a set of vectors of enterprise characteristics at the enterprise level. Let  $\{\boldsymbol{\beta}^j, \boldsymbol{\gamma}^j, \boldsymbol{\delta}^j\}$  denote the associated vectors of coefficients for equation  $j$ , where  $j = 1, 2, 3$ .  $\lambda_t^j$  are time effects for equation  $j$ ,  $\mu_h^j$  are fixed effects at level  $k \in \{h\}$  for equation  $j$ , and  $u_{eht}^j$  are idiosyncratic errors for equation  $j$ .

As individual characteristics we include the gender, age, marital status, and education of the enterprise owner, and the information if the owner is a migrant. As household characteristics we include the number of adults and experience of shocks. And as enterprise characteristics we include the use of credit to expand the business, firm size and months of enterprise operation.<sup>9</sup>

## 5 Regression Results

We report our regression results as follows. First, we test our expectations about the determinants of labor productivity in non-farm enterprises by estimating a Heckman selection model using data from all four countries. Second, we apply panel data methods to refine these results and to study the determinants of labor productivity over time, using data from Nigeria and Uganda.

### 5.1 Heckman Selection Model

In Table 1 we estimate a set of regressions using data from Ethiopia, Malawi, Nigeria, and Uganda. In the upper part of the table (the second stage of the estimation) the dependent variable is the log of labor productivity, and the independent variables entail possible determinants of labor productivity. Individual characteristics are taken from the enterprise owner. In the lower part of the table (the first stage of the estimation) the dependent variables is the binary variable NFE, indicating if a household operates a non-farm enterprise. The independent variables entail possible determinants for the household decision to operate a non-farm enterprise. Individual characteristics are taken from the household head, where we make the assumption that the head decides within the

<sup>9</sup> See Appendix C.2 for summary statistics and a detailed definition of the variables used in the panel data analysis.

household if a household member enters the non-farm business sector. For the selection stage we also make the assumption that the number of adults in a household selects households into this sector, as surplus labor is allocated into enterprises. The reported results in the second stage can then be interpreted as though we observed data for all households in the sample.

In the first stage, the selection stage, we observe that number of adults is significant and positive in all four countries. Households located in a rural area have in all cases a lower likelihood of operating a non-farm enterprise, but the gender of the household head does not seem to determine enterprise operation, with the exception of Nigeria. Age is significant and negative in all regressions, and education, proxied with the ability to read & write, significant and positive, with the exception of Uganda. The variable general access to credit, available in Ethiopia and Malawi, is significant and positive for the regressions of these two countries, suggesting that access to credit significantly increases the likelihood of enterprise operation. Shocks are not significant in most cases with the exception of Malawi, where they increase the likelihood of entering the non-farm sector, possibly indicating enterprise operation due to push factors. Significant and negative effects are reported for the distance variable in Ethiopia and Malawi, and a marginally significant and positive outcome in Uganda.

In the second stage, the outcome stage, we observe that the effect of the variable rural on labor productivity is significant and negative in Malawi and Nigeria, showing that rural enterprises are less productive than urban ones. Also the effect of female is significant and negative in Ethiopia, Malawi, and Nigeria, indicating that female enterprise owners operate less productive businesses. Age does not seem to be associated with labor productivity and is only marginally significant and positive in Uganda. The effect of the ability to read & write is significant and positive for all countries, indicating that more educated enterprise owners also operate more productive businesses. The effect of access to credit is only marginally significant in Malawi and negative, a surprising outcome given that the credit variable has a positive effect on enterprise operation. One explanation could lie in the fact that this variable indicates general access to credit and could be used for any other purpose within the household. The effect of firm size, defined as number of workers per firm, is significant and negative in Ethiopia and Nigeria. While we would expect that larger enterprises are more productive, a possible explanation might lie in the lack of information on the time-use of workers, for example as part-time family workers that contribute little to the enterprise's productivity. The effect of shocks is significant and negative, as we would expect, since exposure to shocks might force households to operate enterprises in less risky, but also lower return types of activity. Finally, the effect of distance has a significant and positive effect in Malawi, but a significant and negative effect in Uganda. While the negative result can be easily explained by access to demand and markets, the positive result is surprising. It could indicate deep rural isolation from competition, and give small, rural enterprises some form of monopolist power.

Summarizing this section, we identify a set of variables that can explain enterprise operation in a first stage, and productivity levels of these enterprises in a second stage. In the first stage, or selection stage, when households decide if they operate an enterprise, the number of adults, the location of the household in a rural or urban area, and the age of the household head are significant factors for all countries. Further education in three

**Table 1:** Heckman Selection Model

	Ethiopia (1)	Malawi (2)	Nigeria (3)	Uganda (4)
<b>(ln) Productivity</b>				
Rural	-0.007 (0.26)	-0.596*** (0.11)	-0.178** (0.07)	0.365 (0.33)
Female	-0.596*** (0.15)	-0.587*** (0.07)	-0.238*** (0.05)	-0.085 (0.16)
Age	0.007 (0.01)	-0.000 (0.00)	-0.001 (0.00)	0.011* (0.01)
Read & Write	0.338** (0.16)	0.321*** (0.08)	0.303*** (0.06)	0.789*** (0.18)
Credit	-0.017 (0.17)	-0.189* (0.10)		
Firm Size	-0.194** (0.08)	0.021 (0.02)	-0.143*** (0.03)	-0.015 (0.03)
Shock	-0.361** (0.15)	-0.412*** (0.08)	-0.131** (0.07)	-0.277 (0.20)
Distance	-0.243 (0.30)	0.526*** (0.17)	-0.261 (0.16)	-1.884*** (0.61)
<b>NFE</b>				
Number of Adults	0.041** (0.02)	0.112*** (0.01)	0.144*** (0.02)	0.066*** (0.01)
Rural	-1.026*** (0.09)	-0.444*** (0.06)	-0.461*** (0.05)	-0.357*** (0.10)
Female	0.110 (0.09)	-0.050 (0.04)	-0.157** (0.07)	-0.052 (0.06)
Age	-0.011*** (0.00)	-0.008*** (0.00)	-0.007*** (0.00)	-0.014*** (0.00)
Read & Write	0.176** (0.08)	0.249*** (0.04)	0.172*** (0.05)	0.042 (0.05)
Credit	0.340*** (0.08)	0.257*** (0.05)		
Shock	0.065 (0.07)	0.084** (0.04)	-0.023 (0.05)	0.067 (0.07)
Distance	-0.357*** (0.13)	-0.314*** (0.09)	-0.125 (0.13)	0.348* (0.20)
Agro-Ecological Zone	Yes	Yes	Yes	Yes
N	3,892	12,496	5,859	2,477
rho	-0.297	-0.611	-0.488	-0.935
sigma	1.577	1.500	1.368	3.099
lambda	-0.469	-0.916	-0.667	-2.899
Prob > chi2	0.107	0.000	0.000	0.000

Standard errors in parentheses; Household weights included; Clustered at the household level.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

out of the four countries studied, and access to credit and distance to the next population center for Ethiopia and Malawi. In the second stage, or outcome stage, the effect of the ability to read & write has a significant and positive effect on labor productivity in all countries. The effects of a household's location, the gender of the enterprise owner, and the experience of shocks are further significant factors in three out of four countries.<sup>10</sup>

<sup>10</sup> See also Table 6 in Appendix D.1 for the same set of heckman selection regressions using rural households only. We do not find major differences using the rural sample only, and conclude that

## 5.2 Panel Data Analysis

As the last part of our analysis we make use of the availability of panel data to further investigate the determinants of enterprises productivity over time. We estimate panel regressions at the enterprise, household, and community level using data from Nigeria and Uganda, where we introduce fixed effects and time fixed effects.

In Table 2 the results of the panel analysis are reported. The variable rural shows that location has a significant and negative effect on labor productivity in Nigeria at the enterprise and household level, but no effect in Uganda at all levels. With regard to the individual characteristics, the effect of gender is, as expected, significant in both countries at all levels. If the enterprise owner is female, productivity is significantly lower compared to a male-owned enterprise, holding other factors constant. The interaction term female x married is significant in Uganda at the enterprise level, and has a negative coefficient, suggesting that women might face a different role with respect to her time-use (e.g. unpaid care and household work). The variable age is significant at the enterprise and household level in Nigeria, and at all levels in Uganda, and suggests that older enterprise owners operate more productive enterprises, which can be possibly explained by experience. The effect of education, proxied by the ability to read & write, does not have an impact in most regressions with the exception of Nigeria at the enterprise level, where it is positive. A surprising outcome that we may explain with the fact that most enterprises are rather small household businesses that do not require high skilled enterprise owners in terms of schooling, and where other abilities that can be learned outside of school might be of higher importance for labor productivity. This outcome is still in contrast with the results of the Heckman selection model and demand further research. Finally, the migration status of the enterprise owner does not report significant effects in Uganda.

In terms of household characteristics, we find that the number of adults within a household is not significant in Nigeria, but positive and significant in Uganda at the enterprise and household level, indicating a higher productivity level if households have more adults to rely upon. The effect of shocks is not significant in Uganda, and marginally significant and positive in Nigeria at the enterprise and household level. This outcome stands again in contrast to the results of the Heckman selection model.

In terms of enterprise characteristics, we report that the effect of firm size has a significant and negative outcome for most regressions. As previously discussed, we assume that this outcome might be explained by the lack of information on the time-use of workers. Having taken out credit to expand the business is positive and significant in most regression results, leading to higher labor productivity, as we would expect based on the literature. One interesting and significant variable that is available for Uganda, is months in operation: the more months per year an enterprise operates its business, the higher its labor productivity, possibly reflecting better capacity utilization.

Summarizing the panel data analysis we find various outcomes that are comparable to the results of the Heckman selection model, for example the effect of the variables rural,

---

both tables report comparable results.



**Table 2:** Panel Data Analysis

(ln) Productivity	Nigeria Enterprise (1)	Nigeria Household (2)	Nigeria Community (3)	Uganda Enterprise (4)	Uganda Household (5)	Uganda Community (6)
Rural	-0.985*** (0.25)	-0.978*** (0.25)	-0.316 (0.52)	-0.108 (0.47)	0.193 (0.48)	-0.341 (0.55)
Female	-0.715*** (0.12)	-0.710*** (0.21)	-0.538 (0.46)	-0.090 (0.25)	-0.267 (0.34)	-0.161 (0.68)
Age	0.042*** (0.01)	0.048*** (0.02)	0.042 (0.04)	0.091*** (0.02)	0.059** (0.03)	0.114 (0.08)
Age <sup>2</sup>	-0.000*** (0.00)	-0.000*** (0.00)	-0.001 (0.00)	-0.001*** (0.00)	-0.001** (0.00)	-0.002* (0.00)
Married	-0.020 (0.10)	-0.202 (0.19)	-0.064 (0.39)	0.665*** (0.21)	-0.015 (0.30)	0.988* (0.55)
Female x Married	-0.044 (0.12)	0.116 (0.24)	-0.246 (0.50)	-0.680*** (0.26)	-0.218 (0.35)	-0.758 (0.75)
Read & Write	0.103* (0.06)	0.091 (0.09)	-0.212 (0.23)	-0.174 (0.13)	-0.196 (0.15)	0.323 (0.42)
Migration				0.135 (0.11)	0.033 (0.12)	-0.171 (0.32)
Number of Adults	-0.030 (0.04)	-0.041 (0.03)	-0.101* (0.06)	0.097*** (0.03)	0.079*** (0.03)	0.084 (0.09)
Shock	0.114* (0.07)	0.105* (0.06)	0.200 (0.16)	-0.057 (0.08)	-0.106 (0.08)	-0.147 (0.26)
Credit	0.246*** (0.07)	0.206** (0.09)	-0.318 (0.27)	0.440*** (0.11)	0.348*** (0.11)	0.643** (0.28)
Firm Size	-0.125*** (0.04)	-0.083** (0.04)	-0.182*** (0.04)	-0.118*** (0.02)	-0.169*** (0.02)	-0.042 (0.07)
Months in Operation				0.061*** (0.02)	0.047*** (0.02)	0.069** (0.03)
Constant	8.583*** (0.25)	8.537*** (0.42)	8.913*** (1.01)	8.321*** (0.64)	9.284*** (0.64)	7.227*** (1.66)
Time FE	YES	YES	YES	YES	YES	YES
FE	YES	YES	YES	YES	YES	YES
Female	0.000	0.000	0.024	0.000	0.005	0.093
Age	0.000	0.007	0.463	0.000	0.045	0.077
Married	0.869	0.495	0.636	0.006	0.407	0.157
N	7,749	5,048	787	4,495	3,619	905

Standard errors in parentheses; Household weights included.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

female, and firm size. However, we also find diverging outcomes as in the case of education and shocks that call for further research. Different to the previous regression model, we use the variable credit to expand the business in the panel data analysis, which has a significant and positive effect on labor productivity. Credit that clearly serves the purpose of expanding the business does show a positive effect on labor productivity, compared to the previously used variable general access to credit.<sup>11</sup>

<sup>11</sup> See also Table 7 in Appendix D.2 for the same set of panel regressions using rural households only. We do not find major differences using the rural sample only, and conclude that both tables report comparable results.

## 6 Summary and Concluding Remarks

In this paper we analyze the determinants of labor productivity in rural African enterprises and provide an exploratory overview of this sector using the World Bank's recent LSMS-ISA data set. Our results confirm a link between a household's motivation to operate an enterprise and its subsequent productivity. In the remaining section, we first summarize our findings, before discussing policy implications in the second part.

In the descriptive statistics we find that location, gender, education, and the experience of shocks report diverging levels of labor productivity. Using kernel density estimates, we show that rural non-farm enterprises are on average less productive than their urban counterparts, confirming a widely held "stylized fact". In terms of location, our results also reveal a more complex and nuanced reality. Using data from Uganda, we identify differences in regional productivity and by distance to the nearest population center. Non-farm enterprises located in regions that have been subject to a history of conflict show lower labor productivity. The closer an enterprise is located to an agglomeration, for example a secondary town, the higher its productivity. We also find that female-owned enterprises are less productive than male-owned enterprises, that literate owners operate more productive enterprises, and that the experience of a shock leads to lower labor productivity.

We further assess labor productivity by estimating two regression models. In the Heckman selection model we find that the ability to read & write, a proxy for education, is associated with higher labor productivity, and that the experience of a shock reduces labor productivity. In the majority of the countries we also find that rural enterprises are less productive, but that distance can have opposing effects on productivity. Using panel data from Nigeria and Uganda we establish that non-farm enterprises in rural areas are less productive in Nigeria, but not in Uganda, and that female-owned enterprises are less productive in both countries. The results also indicate that productivity increases with age, a proxy for experience. The variable credit to expand the business is significant and positive in both countries, suggesting that credit for business purposes increases labor productivity over time. We also identify a positive association between the months of enterprise operation and labor productivity.

Our results confirm a link between a household's motivation to operate a non-farm enterprise and its subsequent labor productivity. Enterprises that experience adverse events and might be operated due to necessity, e.g. due to shocks, are less productive than enterprises that might be operated as a result of the household utilizing an opportunity, e.g. market access, education, or access to finance. Labor productivity is also higher if enterprises operate throughout the year, most likely due to better capacity utilization.

While we recognize that the data reflects a difficult business environment for enterprises in rural Africa, we also conclude that improvements in employment creation and welfare are feasible if well-designed policies are implemented to support non-farm entrepreneurship. We therefore recommend a set of policy implications to improve these business conditions: on the one hand policies that have the potential to increase labor productivity, on the other hand policies that prevent enterprises from stopping enterprise operations. With

regard to the former, we suggest policies that support the enterprise owner or manager, such as education and training, or access to credit to expand business activities. With regard to the latter, we suggest policies that cushion shocks and protect households from negative external events, for example micro-insurance schemes or social protection policies. If policies are considered that tackle the difficulties from both angles (support entrepreneurs and shield households from detrimental events), they could have the potential to boost labor productivity, which in turn is expected to lead to higher survival rates and employment creation in rural Africa.

Our paper provides a first, explorative analysis of labor productivity in rural African enterprises, but calls for more and better research in this area. One of the principal shortcomings of the current analysis is due to data availability. While the LSMS-ISA database offers a rich data set, including a separate section on non-farm entrepreneurship, we face a number of data challenges. First, the differences in the questionnaires among the countries studied only allows for limited comparability, as certain variables or formulation of questions can be found in one, but not in the other questionnaire. Second, our measure of productivity is partial and imprecise. And third, useful variables and information are lacking and therefore limiting us in our analysis. For example the possibility to match the type of business activity with enterprise information, or more detailed information of enterprises that stopped operating in order to conduct survival analysis. More and better data needs to be collected for a more exhaustive and in-depth analysis of the non-farm enterprise sector. Given the need for rural employment, poverty reduction and welfare improvements, faster productivity growth has the potential to facilitate structural transformation in Africa. The benefits of obtaining improved data on rural non-farm enterprises may be worth the effort.

## Acknowledgement

We are particularly grateful to Luc Christiaensen for his inspiration, many stimulating discussions, and useful comments. We also wish to thank Richard Bluhm, Brian Dillon, Jonathan Kaminski, Amparo Palacios-López, Bob Rijkers, Omar Rodríguez-Torres, Alberto Zezza, and other participants at workshops and conferences in Brussels, Lima, Lusaka, Maastricht, Paris and Washington DC. We gratefully acknowledge the comments of an anonymous reviewer. The usual disclaimer applies.

## References

- Alene, A. (2010). Productivity Growth and the Effects of R&D in African Agriculture. *Agricultural Economics*, 41 (3-4):223–238.
- Ali, M. and Peerlings, J. (2011). Value Added of Cluster Membership for Micro Enterprises of the Handloom Sector in Ethiopia. *World Development*, 39 (3):363–374.
- Alsos, G., Carter, S., and Llunggren, E. (2013). Entrepreneurial Families and Households. *ERC Research Paper No.10*.
- Amin, M. (2011). Labor Productivity, Firm-Size and Gender: The Case of Informal Firms in Argentina and Peru. *Enterprise Note No. 22*, The World Bank.
- Arnold, J., Mattoo, A., and Narciso, G. (2006). Services Inputs and Firm Productivity in Sub-Saharan Africa: Evidence from Firm-Level Data. *World Bank Policy Research Working Paper No. 4048*, Washington DC: The World Bank.
- Blattmann, C., Fiala, N., and Martinez, S. (2014). Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda. *Quarterly Journal of Economics*, 129 (2):697–752.
- Block, S. (2010). The Decline and Rise of Agricultural Productivity in Sub-Saharan Africa since 1961. *NBER Working Paper No. 16481*.
- Bloom, N. and Reenen, J. V. (2010). Why do Management Practices differ across Firms and Countries? *Journal of Economic Perspectives*, 24 (1):203–224.
- Bloom, N., Schankerman, M., and Reenen, J. V. (2013). Identifying Technology Spillovers and Product Market Rivalry. *Econometrica*, 81 (4):1347–1393.
- Brueck, T., Naudé, W., and Verwimp, P. (2013). Business under Fire: Entrepreneurship and Violent Conflict in Developing Countries. *Journal of Conflict Resolution*, 57(1):3–19.
- Christiaensen, L., Weerdt, J., and Todo, Y. (2013). Urbanization and Poverty Reduction: the Role of Rural Diversification and Secondary Towns. *Agricultural*, 44(4-5):435–447.
- Dercon, S. (2009). Rural Poverty: Old Challenges in New Contexts. *World Bank Research Observer*, 24(1):1–28.
- Dethier, J., Hirn, M., and Straub, S. (2010). Explaining Enterprise Performance in Developing Countries with Business Climate Survey Data. *World Bank Research Observer*, 26:258–309.
- Dollar, D., Hallword-Driemeier, M., and Mengistae, T. (2005). Investment Climate and Firm Performance in Developing Countries. *Economic Development and Cultural Change*, 54 (1):1–31.
- Eifert, B., Gelb, A., and Ramachandran, V. (2008). The Cost of Doing Business in Africa: Evidence from Enterprise Survey Data. *World Development*, 36 (9):1531–1546.

- Espeland, R. (2007). When Neighbours Become Killers: Ethnic Conflict and Communal Violence in Western Uganda. *CMI Working Paper 2007*, 3.
- Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity and Profitability? *American Economic Review*, 98 (1):394–425.
- Fox, L., Haines, G., Munoz, J., and Thomas, A. (2013). Africa’s Got Work to Do: Employment Prospects in the New Century. *IMF Working Paper No. 13*.
- Fox, L. and Pimhidzai, O. (2013). Household Non-Farm Enterprises and Structural Transformation: Evidence from Uganda. *Paper presented at the UNU-WIDER Conference on Inclusive Growth in Africa, Helsinki*.
- Fox, L. and Sohnesen, T. (2013). Household Enterprises in Mozambique: Key to Poverty Reduction but not on the Development Agenda? *Policy Research Working Paper No. 6570. Washington DC: The World Bank*.
- Frazer, G. (2005). Which Firms Die? A Look at Manufacturing Firm Exit in Ghana. *Economic Development and Cultural Change*, 53 (3):585–617.
- Haggblade, S., Hazell, P., and Reardon, T. (2010). The Rural Non-Farm Economy: Prospects for Growth and Poverty Reduction. *World Development*, 38 (10):1429–1441.
- Janvry, A. D. and Sadoulet, E. (2006). *Poverty, Inequality and Development*, chapter Progress in the Modelling of Rural Households. Springer.
- Kinda, T., Plane, P., and Veganzones-Veroudakis, M. (2011). Firm Productivity and Investment Climate in Developing Countries: How Does the Middle East and North Africa Perform? *CERDI Etudes et Documents*, 26.
- Loening, J., Rijkers, B., and Söderbom, M. (2008). Nonfarm Microenterprise Performance and the Investment Climate: Evidence from Rural Ethiopia. *Policy Research Working Paper No. 4577*, page Washington DC: The World Bank.
- Martin, P., Mayer, J., and Mayneris, F. (2011). Spatial Concentration and Plant-Level Productivity in France. *Journal of Urban Economics*, 69:182–195.
- Moretti, E. (2004). Workers Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions. *American Economic Review*, 94 (3):656–690.
- Nagler, P. and Naudé, W. (2014). Non-Farm Entrepreneurship in Rural Africa: Patterns and Determinants. *IZA Discussion Paper No. 8008*.
- Nichter, S. and Goldmark, L. (2009). Small Firm Growth in Developing Countries. *World Development*, 37 (9):1435–1464.
- Oseni, G. and Winters, P. (2009). Rural Nonfarm Activities and Agricultural Crop Production in Nigeria. *Agricultural Economics*, 40 (2):189–201.
- Owoo, N. and Naudé, W. (2014). Non-Farm Enterprise Performance and Spatial Autocorrelation in Rural Africa: Evidence from Ethiopia and Nigeria. *IZA Discussion Paper No. 8295*.

- Palacios-López, A. and López, R. (2014). Gender Differences in Agricultural Productivity: The Role of Market Imperfections. *Working Paper No. 164061, University of Maryland, Department of Agricultural and Resource Economics.*
- Rijkers, B. and Costa, R. (2012). Gender and Rural Non-Farm Entrepreneurship. *World Development*, 40 (12):2411–2426.
- Rijkers, B., Söderbom, M., and Loening, J. (2010). A Rural-Urban Comparison of Manufacturing Enterprise Performance in Ethiopia. *World Development*, 38 (9):1278–1296.
- Rosenzweig, M. and Binswanger, H. (1993). Wealth, Weather, Risk and the Composition and Profitability of Agricultural Investments. *Economic Journal*, 103(2):56–78.
- Sabates-Wheeler, R., Lind, J., and Hoddinott, J. (2013). Implementing Social Protection in Agro-Pastoralist and Pastoralist Areas: How Local Distribution Structures Moderate PSNP Outcomes in Ethiopia. *World Development*, 50:1–12.
- Saliola, F. and Seker, M. (2011). Total Factor Productivity across the Developing World. *Enterprise Note No. 23*, The World Bank.
- Singh, I., Squire, L., and Strauss, J., editors (1986). *Agricultural Household Models: Extensions, Applications and Policy*. Johns Hopkins University Press.
- Söderbom, M. and Teal, F. (2004). Size and Efficiency in African Manufacturing Firms: Evidence from Firm-Level Panel Data. *Journal of Development Economics*, 73 (1):369–394.
- Söderbom, M., Teal, F., and Harding, A. (2006). The Determinants of Survival among African Manufacturing Firms. *Economic Development and Cultural Change*, 54 (3):533–555.
- Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature*, 49 (2):326–365.
- Taylor, J. and Adelman, I. (2003). Agricultural Household Models: Genesis, Evolution, and Extensions. *Review of Economics of the Household*, 1(1):33–58.
- van Biesebroeck, J. (2005). Exporting Raises Productivity in Sub-Saharan Manufacturing Firms. *Journal of International Economics*, 67 (2):373–391.
- Wennberg, K. and Lindqvist, G. (2010). The Effect of Clusters on the Survival and Performance of New Firms. *Small Business Economics*, 34 (3):221–241.

# A Appendix: Data Manipulation

## All Estimates

We replace the age of the household head or enterprise owner with the median age in case the age variable is missing, below 15, or above 99. This happens however only for very few observations.

## Panel Data Analysis

Households are visited and interviewed during the first survey round, and households operating a non-farm enterprise are registered in the respective section. Due to the panel design, the same households are visited again in the consecutive survey rounds. If a household of the previous round is no longer operating an enterprise, the household is still registered in the respective section, but does not fill out the questions on non-farm entrepreneurship. Therefore it is possible to identify the enterprises that were discontinued since the last survey round if the enterprises can clearly be tracked over the waves. If not, it is possible to know the number of enterprises that were discontinued. In Uganda, various “empty” observations in follow-up surveys do not have observations in the previous survey round(s), and we decide to exclude these observations, as it is not clear what kind of information they present, if they are discontinued enterprises or false observations. We delete 82 out of 7,985 observations, resulting in a final database of 7,903 enterprises for the panel data analysis. We do not encounter this issue in Nigeria.

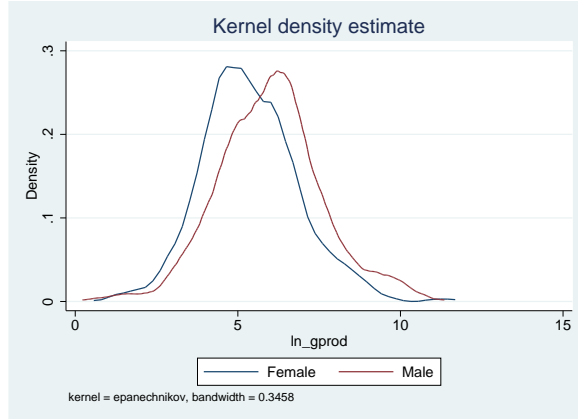
Further data manipulation includes the following: prices are inflation deflated to prices of the first survey round of the respective country, using the IMF’s inflation index of average consumer prices (percentage change). Furthermore prices are “winsorized” at 1 percent. We also “winsorize” the number of employees at 1 percent. We further replace a few outlier values, for example if the months in operation exceeded 12 months.



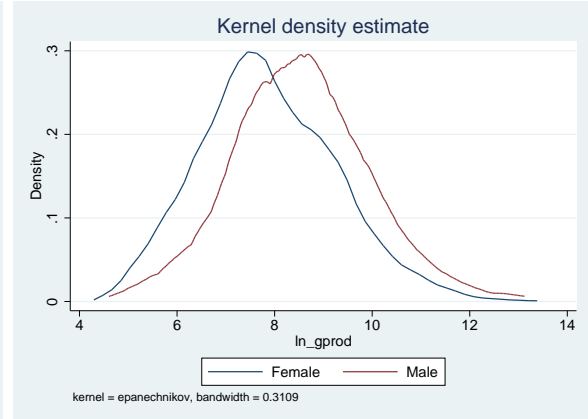
## B Appendix: Productivity Dispersal

Figure 3: Productivity Dispersal by Gender

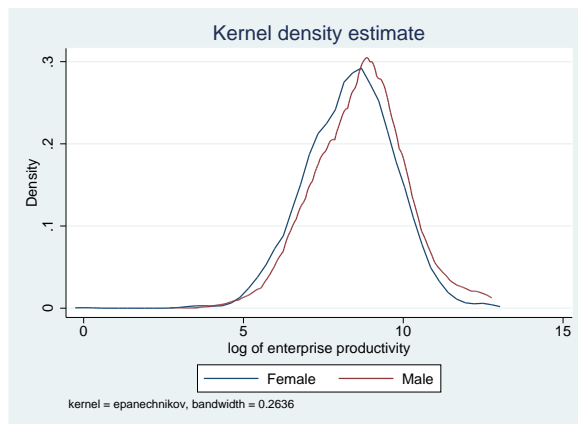
(a) Ethiopia



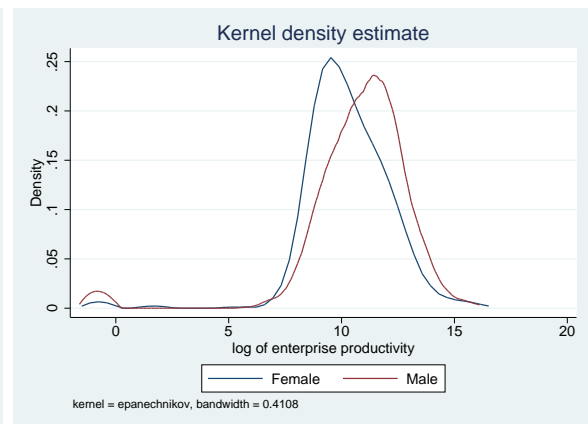
(b) Malawi



(c) Nigeria

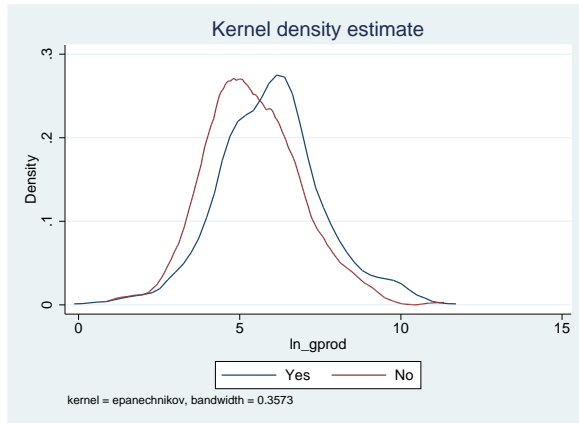


(d) Uganda

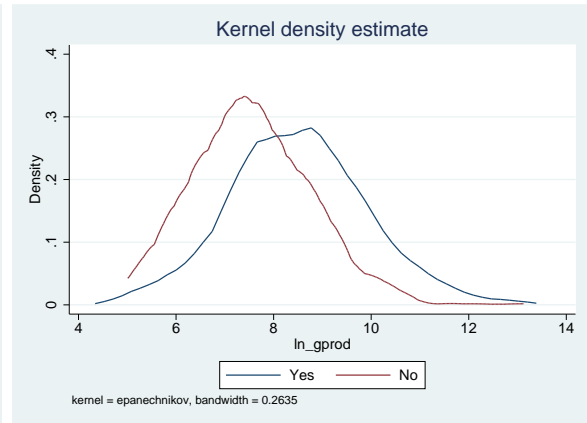


**Figure 4:** Productivity Dispersal by Ability to Read & Write

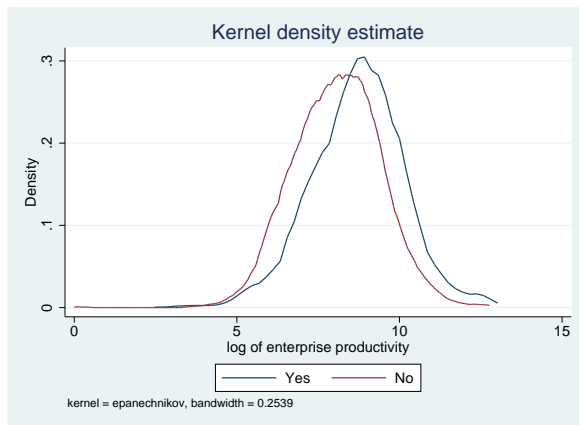
**(a)** Ethiopia



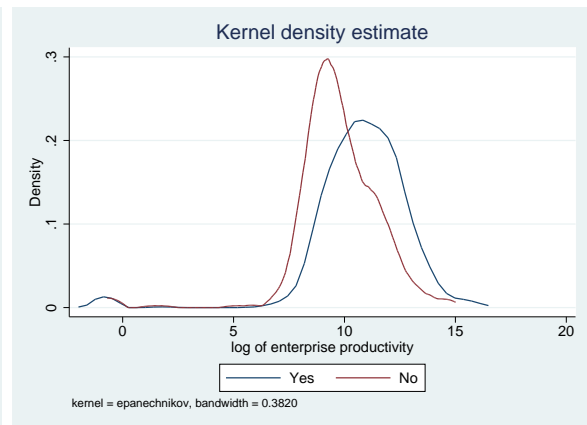
**(b)** Malawi



**(c)** Nigeria

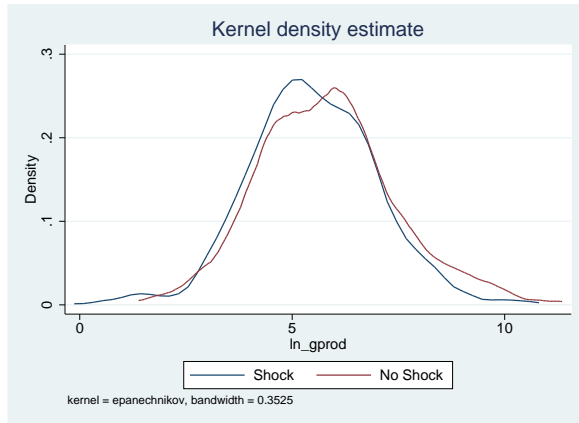


**(d)** Uganda

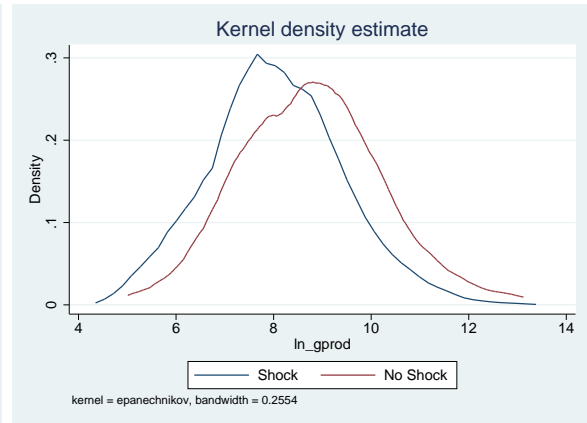


**Figure 5: Productivity Dispersal by Shock Experience**

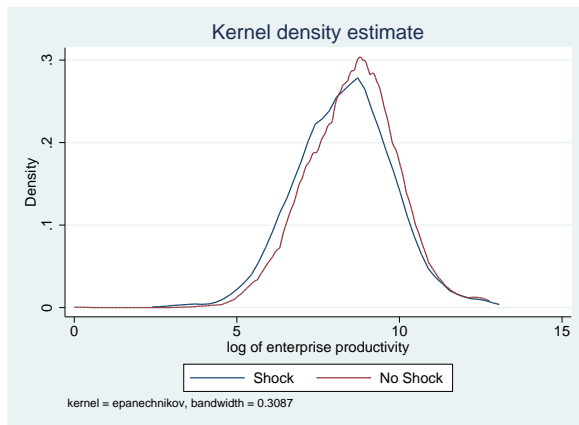
**(a) Ethiopia**



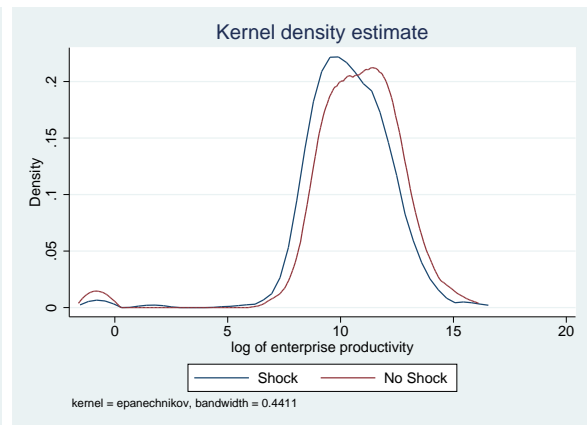
**(b) Malawi**



**(c) Nigeria**



**(d) Uganda**



## C Appendix: Summary Statistics

### C.1 Heckman Selection Model

**Table 3:** Summary Statistics - Heckman Selection Model

	Ethiopia			Malawi		
	Obs	Mean	St Dev	Obs	Mean	St Dev
<b>Individual Characteristics</b>						
Household Head						
Female	4,150	0.25	0.43	12,506	0.24	0.43
Age	4,146	43.96	15.64	12,506	42.08	16.07
Read & Write	4,131	0.41	0.49	12,506	0.68	0.47
Enterprise Owner						
Female	1,344	0.51	0.50	2,802	0.38	0.48
Age	1,342	35.76	12.60	2,802	37.06	12.22
Read & Write	1,336	0.48	0.50	2,802	0.77	0.42
<b>Household Characteristics</b>						
Number of Adults	4,159	4.13	2.16	12,506	2.47	1.22
Rural	4,159	0.86	0.34	12,506	0.81	0.39
Credit	4,087	0.25	0.42	12,506	0.13	0.34
Shock	4,159	0.47	0.50	12,505	0.69	0.46
Distance	4,159	0.41	0.34	12,506	0.34	0.23
<b>Enterprise Characteristics</b>						
Firm Size	1,470	2.25	1.86	2,809	2.59	1.73

**Table 4:** Summary Statistics - Heckman Selection Model

	<b>Nigeria</b>			<b>Uganda</b>		
	Obs	Mean	St Dev	Obs	Mean	St Dev
<b>Individual Characteristics</b>						
<b>Household Head</b>						
Female	6,555	0.13	0.33	3,508	0.29	0.46
Age	6,537	49.36	14.81	3,506	45.68	14.76
Read & Write	6,543	0.64	0.48	3,395	0.66	0.47
<b>Enterprise Owner</b>						
Female	4,364	0.57	0.50	1,909	0.52	0.50
Age	4,358	40.86	13.84	1,907	39.90	12.50
Read & Write	4,354	0.63	0.48	1,895	0.74	0.44
<b>Household Characteristics</b>						
Number of Adults	6,555	3.35	1.89	3,508	4.10	2.51
Rural	6,558	0.66	0.47	3,511	0.76	0.43
Shock	6,558	0.31	0.46	3,444	0.46	0.50
Distance	6,558	0.20	0.20	3,102	0.22	0.19
<b>Enterprise Characteristics</b>						
Firm Size	4,318	2.24	1.57	1,525	3.14	1.89

### Definition of Variables for the Heckman Selection Model

Female	1 if female.
Age	In Years.
Read & Write	1 if individual can read and write in any language.
Number of Adults	Number of adults age 15 or older in the household.
Rural	1 if household is located in a rural area.
Credit	1 if household has access to credit. This variable is defined as general access to credit, and does not further specify for which purpose the household used it.
Shock	1 if household experienced a shock. Shocks can be idiosyncratic (e.g. death or illness of a household member), related to prices (e.g. increase in the price level of certain goods and services), related to agriculture (e.g. droughts or floods), or other types of shocks (not further specified in the questionnaire).
Distance	Defined as distance to next population center of 20,000 or more inhabitants, in 100's of km.
Firm Size	Number of workers in the enterprise.
Agro-Ecological Zones	Tropic-warm or tropic-cool, combined with different precipitation levels: arid, semi arid, sub humid, humid.

## C.2 Panel Data Analysis

**Table 5:** Summary Statistics - Panel Data Analysis

	Nigeria			Uganda		
	Obs	Mean	St Dev	Obs	Mean	St Dev
<b>Individual Characteristics</b>						
Female	8,715	0.56	0.50	4,686	0.49	0.50
Age	8,711	2.96	0.99	4,684	2.83	0.98
Read & Write	8,687	0.64	0.48	4,530	0.75	0.43
Migration				4,796	0.64	0.48
<b>Household Characteristics</b>						
Rural	8,715	0.62	0.49	4,795	0.71	0.45
Number of Adults	8,715	3.90	2.15	4,686	3.71	2.16
Shock	8,715	0.24	0.43	4,796	0.48	0.50
<b>Enterprise Characteristics</b>						
Credit	8,666	0.08	0.26	4,785	0.12	0.33
Firm Size	8,715	2.21	2.06	4,796	4.04	2.28
Months in Operation				4,779	9.59	3.49

### Definition of Variables for the Panel Data Analysis

Female	1 if female.
Age	In Years.
Read & Write	1 if individual can read and write in any language.
Migration	1 if individual migrated. It indicates if the individual moved from another district or country to the current place of residence.
Rural	1 if household is located in a rural area.
Number of Adults	Number of adults age 15 or older in the household.
Shock	1 if household experienced a shock. Shocks can be idiosyncratic (e.g. death or illness of a household member), related to prices (e.g. increase in the price level of certain goods and services), related to agriculture (e.g. droughts or floods), or other types of shocks (not further specified in the questionnaire).
Credit	1 if household has taken out credit to expand the business. The use of credit is clearly related to the purpose of operating and expanding the business, and not used for other purposes within the household.
Firm Size	Number of workers in the enterprise.
Months in Operation	Number of months the enterprise was in operation during the past 12 months (max. 12).

## D Appendix: Regressions - Rural Samples

### D.1 Heckman Selection Model

**Table 6:** Heckman Selection Model - Rural

	Ethiopia (1)	Malawi (2)	Nigeria (3)	Uganda (4)
<b>(ln) Productivity</b>				
Female	-0.593*** (0.15)	-0.614*** (0.07)	-0.333*** (0.07)	-0.137 (0.17)
Age	0.007 (0.01)	-0.003 (0.00)	-0.004 (0.00)	0.013** (0.01)
Read & Write	0.344** (0.16)	0.302*** (0.08)	0.178** (0.08)	0.766*** (0.20)
Credit	-0.005 (0.18)	-0.227** (0.09)		
Firm Size	-0.204** (0.09)	-0.021 (0.03)	-0.179*** (0.03)	-0.025 (0.04)
Shock	-0.363** (0.15)	-0.488*** (0.09)	-0.254*** (0.08)	-0.405* (0.22)
Distance	-0.272 (0.32)	0.510*** (0.18)	-0.263 (0.18)	-1.485*** (0.54)
<b>NFE</b>				
Number of Adults	0.039* (0.02)	0.094*** (0.02)	0.153*** (0.02)	0.071*** (0.01)
Female	0.106 (0.09)	-0.069 (0.05)	-0.163** (0.08)	-0.047 (0.07)
Age	-0.011*** (0.00)	-0.007*** (0.00)	-0.005*** (0.00)	-0.016*** (0.00)
Read & Write	0.179** (0.08)	0.260*** (0.04)	0.263*** (0.06)	0.045 (0.05)
Credit	0.338*** (0.08)	0.320*** (0.05)		
Shock	0.069 (0.07)	0.084** (0.04)	-0.047 (0.06)	0.102 (0.07)
Distance	-0.373*** (0.14)	-0.401*** (0.10)	-0.124 (0.14)	0.218 (0.21)
Agro-Ecological Zone	Yes	Yes	Yes	Yes
N	3,376	10,146	3,891	1,940
rho	-0.277	-0.554	-0.494	-0.916
sigma	1.577	1.407	1.423	2.897
lambda	-0.437	-0.780	-0.703	-2.655
Prob > chi2	0.149	0.000	0.000	0.000

Standard errors in parentheses; Household weights included; Clustered at the household level.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

## D.2 Panel Data Analysis

**Table 7:** Panel Data Analysis - Rural

(ln) Productivity	Nigeria	Nigeria	Nigeria	Uganda	Uganda	Uganda
	Enterprise	Household	Community	Enterprise	Household	Community
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.807*** (0.17)	-0.837*** (0.30)	-0.228 (0.56)	-0.196 (0.32)	-0.217 (0.40)	-0.205 (0.80)
Age	0.034*** (0.01)	0.019 (0.02)	0.036 (0.05)	0.109*** (0.03)	0.065** (0.03)	0.154* (0.09)
Age <sup>2</sup>	-0.000*** (0.00)	-0.000 (0.00)	-0.001 (0.00)	-0.001*** (0.00)	-0.001* (0.00)	-0.002* (0.00)
Married	-0.020 (0.13)	-0.062 (0.23)	0.193 (0.46)	0.453* (0.24)	0.151 (0.30)	1.448** (0.66)
Female x Married	0.014 (0.19)	0.160 (0.32)	-0.489 (0.62)	-0.505 (0.33)	-0.219 (0.41)	-0.890 (0.88)
Read & Write	0.029 (0.07)	-0.061 (0.10)	-0.270 (0.25)	-0.095 (0.16)	-0.133 (0.18)	0.357 (0.51)
Migration				0.106 (0.12)	0.106 (0.14)	-0.421 (0.41)
Number of Adults	-0.048 (0.04)	-0.045 (0.04)	-0.045 (0.07)	0.094** (0.04)	0.057 (0.04)	0.042 (0.10)
Shock	0.106 (0.09)	0.072 (0.07)	0.073 (0.18)	-0.006 (0.09)	-0.068 (0.10)	-0.306 (0.32)
Credit	0.326*** (0.09)	0.211* (0.12)	-0.229 (0.30)	0.426*** (0.13)	0.394*** (0.14)	0.392 (0.36)
Firm Size	-0.201*** (0.04)	-0.285*** (0.04)	-0.185*** (0.06)	-0.129*** (0.03)	-0.180*** (0.03)	-0.125 (0.10)
Months in Operation				0.035* (0.02)	0.037** (0.02)	0.064 (0.04)
Constant	8.177*** (0.26)	8.828*** (0.48)	8.347*** (1.09)	7.914*** (0.73)	8.991*** (0.67)	6.168*** (1.93)
Time FE	YES	YES	YES	YES	YES	YES
FE	YES	YES	YES	YES	YES	YES
Female	0.000	0.000	0.100	0.000	0.065	0.062
Age	0.001	0.583	0.141	0.001	0.099	0.134
Married	0.988	0.881	0.705	0.172	0.864	0.056
N	4,805	3,116	558	3,228	2,650	658

Standard errors in parentheses; Household weights included.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$