Understanding the Role of Rural Non-Farm Enterprises in Africa’s Economic Transformation: Evidence from Tanzania

XINSHEN DIAO*, EDUARDO MAGALHAES** & MARGARET MCMILLAN*†

*Development Strategy and Governance Division, International Food Policy Research Institute, Washington, DC, USA,
**Independent Consultant, †Department of Economics, Tufts University, Boston, MA, USA

ABSTRACT Tanzania’s recent growth boom has been accompanied by a threefold increase in the share of the rural labour force working in nonfarm employment. Although households with nonfarm enterprises are less likely to be poor, a substantial fraction of these households fall below the poverty line. Heterogeneity in the labour productivity of rural nonfarm businesses calls for a two-pronged strategy for rural transformation. Relatively unproductive enterprises may be part of a poverty reduction strategy but should not be expected to contribute to employment and labour productivity growth. Failure to account for this heterogeneity is likely to lead to disappointing outcomes.

1. Introduction

Since the beginning of the twenty-first century, Tanzania’s economy has grown more rapidly than at any other point in recent history. Between 2000 and 2015, the average annual GDP growth rate was 6.8 per cent and the average annual labour productivity growth rate was more than 4 per cent. Between 2002 and 2012, more than three quarters of this labour productivity growth was accounted for by structural change; the remainder of the growth is largely attributable to within sector productivity growth in agriculture. The growth attributable to structural change is almost entirely explained by a rapid decline in the agricultural employment share and an increase in the non-agricultural private sector employment share (Diao, Kweka, McMillan, & Qureshi, 2017).

Despite these changes, Tanzania remains heavily rural; between 2002 and 2012, the share of the population living in rural areas declined by only 6.5 percentage points: from 76.9 per cent to 70.4 per cent (Table 1). Living in rural areas is traditionally associated with farming. However, the census and household survey data also shows that between 2002 and 2012, the share of the rural population engaged in agricultural activities decreased by almost 14 percentage points (Table 1). The data also show that growth in rural nonfarm employment has been very rapid at between 11.1 per cent and 13.5 per cent per annum depending on the data source (Table 1). Thus, while the agricultural sector still employs the majority of the rural population in Tanzania, the rural nonfarm economy is becoming increasingly important.

The purpose of this paper is threefold. The first is to describe the characteristics of the households that make up the rural nonfarm sector in Tanzania. The second is to describe the businesses run by households in the rural nonfarm sector; we also describe characteristics of the owners of these businesses and their self-reported motivations for running a business. The final purpose of this paper is to assess the productivity of rural nonfarm businesses. This last exercise is meant to inform the following question: does the rural nonfarm sector have the potential to contribute to long-run
productivity growth in Tanzania. The importance of this question is highlighted in recent work by Diao, McMillan, and Rodrik (2017) who show that: (i) the pattern of growth in Tanzania is common across Africa and; (ii) without labour productivity growth in nonfarm enterprises, labour productivity growth in Africa is likely to stall.

Using Tanzania’s 2012 Household Budget Survey (HBS) – we classify rural households into three groups: (i) households uniquely engaged in farming; (ii) households uniquely engaged in the nonfarm sector and; (iii) mixed households. Our analysis shows that more than 10 per cent of Tanzanian rural households engage only in the nonfarm economy. The heads of these households tend to be younger and better educated than the heads of households in the two comparison groups. Gender of the household head does not seem to affect the likelihood of being a nonfarm household. Households in

| Year  | Share of rural population (percentage) | Rural areas | National total |
|-------|----------------------------------------|-------------|----------------|
| 2002  | 76.9                                   |             |                |
| 2012  | 70.4                                   |             |                |

| Year              | Share of agricultural employment in total employment (percentage) | Rural areas | National total |
|-------------------|---------------------------------------------------------------|-------------|----------------|
| 2000/01 (HBS)     | 91.8                                                           |             | 82.6           |
| 2002 (Census)     | 93.2                                                           |             | 81.1           |
| 2006 (ILFS)       |                                                               |             | 76.5           |
| 2011/2012 (HBS)   | 77.4                                                           |             | 63.8           |
| 2012 (Census)     | 79.5                                                           |             | 65.8           |
| 2014 (ILFS)       |                                                               |             | 66.9           |

Annualised growth rate in population, total employment and employment in agriculture and non-agriculture (percentage)

| Year              | Population (2002–2012, Census) | Rural areas | National total |
|-------------------|---------------------------------|-------------|----------------|
|                   | 1.8                             |             | 2.7            |

| Year              | Total employment | Rural areas | National total |
|-------------------|------------------|-------------|----------------|
| 2000–2011 (HBS)   | 1.3               |             | 2.4            |
| 2002–2012 (Census)| 1.6               |             | 2.5            |
| 2006–2014 (ILFS)  | 0.3               |             | 2.4            |

| Year              | Employment in agriculture | Rural areas | National total |
|-------------------|---------------------------|-------------|----------------|
| 2000–2011 (HBS)   | −0.3                       |             | 0.1            |
| 2002–2012 (Census)| 0.1                       |             | 0.4            |
| 2006–2014 (ILFS)  |                           |             | 0.7            |

| Year              | Employment in non-agriculture | Rural areas | National total |
|-------------------|------------------------------|-------------|----------------|
| 2000–2011 (HBS)   | 11.1                         |             | 9.4            |
| 2002–2012 (Census)| 13.5                        |             | 8.8            |
| 2006–2014 (ILFS)  |                             |             | 6.8            |

Notes: For the Census employment, data for current employees aged 10 years old and above is used. Agricultural employment is based on the industry classification. For the Household Budget Survey (HBS), employees are for aged 10 years old and above. The definition of employment differs between the two rounds of HBS. In HBS 2000/2001, ‘unpaid family helper’ counts for 7.8 per cent of total employment and is all considered as non-agricultural employment. This causes two problems: (1) total employment of HBS 2000/2001 is more than that of Census 2002, which was conducted two years later; and (2) share of non-agricultural employment in HBS 2000/2001 is too high. Moreover, given that population calculated from HBS 2000/2001 is lower than that in Census 2002, the ratio of employment to population in HBS 2000/2001 becomes too high compared with that of Census 2002. For these reasons, we decided to exclude ‘unpaid family helpers’ from total employment in HBS 2000/2001. After this adjustment, the ratio of total employment to population is similar between HBS 2000/2001 and Census 2002, as well as the shares of agricultural and non-agricultural employment. Employees in the Integrated Labor Force Survey (ILFS) are for age 15 years old and above, and the data for rural employment is not available in the survey report published by Tanzania National Bureau of Statistics (NBS) (2015).

Source: Most numbers in the table are calculated by authors using the micro data downloaded from the NBS website as well as from official government documents. For Census 2012 and ILFS 2006 and 2014, the micro data is not available, and the calculation for these surveys is solely based on official government documents published by NBS (NBS, 2006, 2011a, b, 2014a, b, 2015). The micro data for Census 2002 is downloaded from IPUMS (https://usa.ipums.org/usa/).
communities with access to daily public transportation are found to have a higher probability of participating in rural nonfarm activities.

In order to analyse business in the rural nonfarm sector, the Tanzania National Baseline Survey of Micro, Small and Medium Enterprise (MSME) has been employed. The survey has been conducted by the Tanzania government, and it was the first nationally representative survey of MSME. It is also one of the only comprehensive nationally representative surveys of MSME types in sub-Saharan Africa. It covers roughly three million businesses and five million people; around 75 per cent of these businesses are in rural areas. 20 per cent of rural nonfarm businesses operate in the manufacturing sector while the remainder operate in various service sectors.

Our firm level analysis begins with a description of the characteristics of the owners of these businesses; roughly half say they would not quit their business for a full-time salaried position. We then document the enormous degree of productive heterogeneity in both rural and urban enterprises. Using the information about the productive heterogeneity of the rural nonfarm enterprises, we identify a group of small firms which appear to have the potential to contribute significantly to rural transformation. A probit analysis is conducted to better understand which readily observable characteristics of these firms and their owners are correlated with labour productivity. The analysis is an attempt to search for firm and/or owner attributes that might be used by policy-makers for targeting services to a select group of firms.

This research contributes to a large and growing literature on the rural nonfarm economy in Africa. One strand of this literature focuses on farm/nonfarm linkages and the estimation of multipliers (see for example Haggblade, Hazell, & Brown, 1989). A second strand of this literature focuses on rural-urban linkages and stresses the importance of reducing barriers to the movement of labour and products from rural to urban areas (see for example, De Brauw, Mueller, & Lee, 2014 on migration; and Gollin & Rogerson, 2009 on transportation costs). A third strand of this literature studies the effects of farm/nonfarm linkages and rural/urban linkages simultaneously; (see for example, Haggblade, Hazell, & Reardon, 2010). Unlike these papers, our work focuses on the productivity of rural nonfarm businesses and their potential to contribute to economy-wide labour productivity growth.

This research also contributes to a large body of work on SMEs in developing countries. Numerous studies in this field focus on documenting and analysing the heterogeneity of SMEs in both rural and urban areas (Banerjee, Breza, Duflo, & Kinnan, 2015; Bezu & Barrett, 2012; Fafchamps, McKenzie, Quinn, & Woodruff, 2014; Grimm, Krüger, & Lay, 2011; La Porta & Shleifer, 2011, 2014; McKenzie, 2015; McKenzie & Woodruff, 2008; Nagler & Naudé, 2017; Nordman, Rakotomanana, & Roubaud, 2016; Rijkers & Costa, 2012; Rijkers, Soderbom, & Loening, 2010). Some other studies in the SME literature focus on assessing employment growth in SMEs (McPherson, 1996; Mead, 1994; Mead & Liedholm, 1998), while others focus on the factors that constrain SME development (Beck, 2007; Gibson & Olivia, 2010; Jin & Deininger, 2009; Raj & Sen, 2013).

Most of this previous work relies on small samples and where samples are nationally representative, labour productivity is not examined since the nationally representative surveys do not include operating data. This point was made recently by Li and Rama (2015) in an article published in the World Bank Research Observer. According to Li and Rama (2015), in most African countries enterprise censuses or surveys generally cover formally registered firms, therefore excluding the vast majority of micro- and small-enterprises, which are typically informal (Li & Rama, 2015). The surveys that do cover SMEs are often small in sample size. It follows that the role of small, largely informal firms in the growth and development of poor economies is likely not well known using data from such small sized surveys, given the large heterogeneity among SMEs in developing countries. In addition, analyses of the rural nonfarm economy typically do not compare the productivity of rural enterprises to the productivity of urban enterprises.4

The remainder of this paper is organised as follows. Section 2 demonstrates the growing importance of the rural nonfarm economy in the context of the broader Tanzanian economy. Section 3 describes the data and methods used in the analyses. Section 4 focuses on the household level analysis and examines the characteristics of households with and without rural nonfarm activities. Section 5 turns to the firm data and describes the characteristics of rural entrepreneurs and their businesses. This section
also identifies a group of firms – the ‘in-between’ firms – with the potential to contribute to rural transformation. Section 6 studies the characteristics of the ‘in-between’ firms to better understand how policy-makers might target these firms. Section 7 concludes with a summary of the main points and a brief discussion of policy implications.

2. The role of rural nonfarm enterprises in Tanzania’s economy

Three nationally representative surveys are used to describe the changes in shares of rural population and rural employment. These are the 2000/2001 and 2011/2012 rounds of the Household Budget Surveys (HBS), the 2002 and 2012 Population Censuses, and the 2006 and 2014 Integrated Labor Force Surveys (ILFS). According to the two rounds of the census, which capture the change in population structure, the share of Tanzania’s population living in rural areas declined from 76.9 per cent in 2002 to 70.4 per cent in 2012 (Table 1, first panel). Thus, Tanzania is urbanising but still heavily rural.

The second panel of Table 1 contains agricultural employment shares from the three different surveys in their different rounds. While different surveys cannot be directly compared because of differences in the definition for agricultural employment, there is a clear trend in which the share of agricultural employment declines much more rapidly than the share of rural population. Between the two rounds of the Census (2002–2012), the agricultural share of employment declined by 15.3 percentage points nationally and 13.7 percentage points in rural areas. The two rounds of HBS conducted in 2000/2001 and 2011/2012 and the two rounds of ILFS conducted in 2006 and 2014 all show a similar trend (Table 1). Comparing these large declines with the modest changes in the share of rural population discussed above indicates that growth in rural nonfarm employment outpaced growth in agricultural employment over this period.

Indeed, the annualised employment growth rates in the recent 15 years presented in the third panel of Table 1 clearly indicate this pattern. These employment growth rates are computed from each of the three surveys’ two most recent rounds. In general, employment in agriculture has been growing at a slower pace than total employment nationwide and in rural areas. According to the census data, the growth rate of agricultural employment is almost zero between 2002 and 2012 in rural areas. By contrast, growth in rural non-agricultural employment has been in the double-digit range growth rate over this period (the bottom of Table 1).

To help better understand the changing structure of Tanzania’s employment, Table 2 displays the structure of net increases in employment across different economic sectors between 2002 and 2012 for total employment and formal and informal employment. While the agricultural sector still accounts for two-thirds of total employment in Tanzania, as shown in Table 1, it has played a relatively minor role in the net job increase as shown in Table 2. In fact, almost 90 per cent of the net increase in jobs occurred in the non-agricultural sector over the period of 2002–2012. Considering that agricultural employment made up more than 80 per cent of total employment nationwide in 2002 (Table 1), this rapid non-agricultural employment growth is remarkable.

As is evident from Table 2, about 73 per cent of the net increase in total employment has taken place in the informal non-agricultural sector, accounting for 88 per cent of the increase in private non-agricultural employment. We do not have access to detailed employment data disaggregated by rural and urban among different sectors. However, we know from Table 1 that nonfarm employment increased significantly in rural areas. If one assumes that formal non-agricultural employment is more likely to take place in urban than in rural areas, which is a realistic assumption, then the 88 per cent of increased private sector’s non-agricultural employment created by the informal sector nationwide can be taken as the lower bound for informal employment growth in rural areas. Section 4 of this paper will use the MSME data to further investigate the nature of this rural nonfarm employment.
3. Data and methods

The two datasets used for the analyses are briefly discussed in this section, followed by a description of the methodologies employed to analyse two distinct research questions. We provide a more detailed description of the HBS data in the Appendix since the data are fairly standard and have been used frequently. We devote more space here to the description of the MSME data which is unlikely to be familiar to readers both because of its’ limited use and because these kinds of nationally representative firm level surveys of mostly informal firms are much rarer. First, what are the characteristics of households which participate in the rural nonfarm economy compared to other types of households? And second, can the firm level data be used to assess the characteristics and potential of rural enterprises in comparison with those in urban areas?

3.1. The HBS data

The 2011/2012 Household Budget Survey (HBS) is a nationally representative survey, which is designed to provide estimates of household income and expenditures for poverty assessments, similar to how the Living Standard Measurement Surveys (LSMSs) are conducted routinely in many other countries. The 2011/2012 HBS surveyed 4130 rural households and 6056 rural households, and is also representative at three geographic locations — rural, Dar es Salaam and other urban. Similar to a standard LSMS, the HBS has an occupational module that provides information for all household members’ primary employment by industries (including farming). This is the module used for the rural employment analysis in this paper. The survey also asked whether the households have their home enterprises, a question also used in the discussion in Section 3. However, the HBS only covers mainland Tanzania and excludes Zanzibar. Like the MSME survey to be discussed later, the sampling framework used to conduct the HBS survey is based on the 2002 Census, which could possibly oversample rural households given that, as discussed in Section 1, the 2012 Census has shown a decline of 6.5 percentage points in the share of rural population from the 2002 Census. A set of summary statistics based on the 2011/2012 HBS for the variables used in our analysis is presented in Appendix Table A1.
3.2. The MSME data

As mentioned in Section 1, the Micro, Small, and Medium Sized Enterprise (MSME) survey is Tanzania’s first nationally representative survey of small businesses. The data was collected during interviews with 6134 small business owners identified in a three-step sampling process. In the first step, a sample of 640 representative enumeration areas were selected. A complete listing of all households was carried out to identify households that currently owned and ran small businesses or had recently closed businesses. In the second step, about nine to 12 households with currently operating businesses (and two to three households with closed businesses) were selected. Finally, if more than one member in a selected household owned and ran a small business, a Kish Grid was applied to select the interviewee. There are three questionnaires for the survey. The main questionnaire, which is used for this study, includes 192 questions on 20 topics that are asked to owners of currently operating enterprises. Based on the enterprise’s main activities, main products and services, all enterprises in the survey were assigned to an industry according to the International Standard for Industrial Classification (ISIC). There are 80 unique ISIC industries, of which many have few sampled firms. Thus, in this study, the 80 industries are aggregated into 24 subsectors, six of which are in the manufacturing sector and the rest in service sectors.

A set of summary statistics of the MSME survey data is reported in Table 3 for the three areas separately: rural, other urban area and Dar es Salaam. Among the 6134 sampled firms, a total of 5609 firms have all the information required for the analysis. Based on the information that is available, there is no reason to believe that the firms with missing information are significantly different from the rest of the sample; for example, they are dispersed across regions and firm size.

As shown in the first row of Table 3, most MSMEs are extremely small: mean employment is 1.5 in rural areas, 1.65 in urban areas outside Dar es Salaam and 1.7 in Dar es Salaam. The very small size of MSMEs is at least in part possibly due to a sample selection bias. The sampling framework is household based rather than enterprise based, and is based on the 2002 Population and Housing Census. This could possibly lead to oversampling rural households given that, as discussed in Section 1, the 2012 Census has shown a decline of 6.5 percentage points in the share of rural population from the 2002 Census. The household-based sampling means that the survey probably under-sampled businesses outside households, which in practice translates into under-sampling relatively larger sized firms. The fact that more than 40 per cent of firms with five or more employees in the survey are sampled in rural areas might support this concern, as it is known that most larger sized firms are often in urban areas instead of rural areas. Among the 6134 enterprises sampled in the survey, there are only 96 firms with five or more employees. While for the largest firm in the sample there are 80 employees, the second and third largest ones have employees of 34 and 33 respectively. In fact, there are only four firms in the survey with employees more than 20. While the larger sized firms are usually assigned larger weights than the smaller ones in the data, this potential issue of possibly under-sampling larger urban firms is unlikely to be fully corrected by the assigned sample weights given that the sampling weights are derived from the relationship between listed households and households currently operating with businesses in selected representative enumeration areas.

As expected, few small firms are registered with Tanzania’s Business Registration and Licensing Agency (BRELA) and there is little difference between rural and urban firms in this regard. In contrast, more urban enterprises (8%) have tax identification numbers than rural enterprises (3%). While the MSME survey is a household based survey, 49 per cent of rural firms report that their businesses actually operate out of their homes and the number in urban areas is almost identical at 47 per cent. Therefore, the shares of nonfarm enterprises reported in Table 4 calculated from HBS 2011/2012 data could significantly underestimate the importance of MSMEs in rural areas, given that HBS captures only businesses run out of the home.

Table 3 also reports average monthly value-added and average monthly sales per firm. Firms in the MSME database report sales on a monthly basis and also provide their own judgement on whether a particular month is a good, bad or normal month. By taking the possible seasonality into account, value added is then computed as the firm’s average monthly sales minus the firms’ average monthly
| Names of variables | Value unit or range | Observations Mean | S.D. | Observations Mean | S.D. | Observations Mean | S.D. |
|--------------------|--------------------|-------------------|-----|-------------------|-----|-------------------|-----|
| Number of employees per firm | Person | 4,160 | 1.40 | 1.01 | 1093 | 1.64 | 1.94 | 353 | 1.74 | 1.52 |
| Number of full-time employees per firm | Person | 4,150 | 1.23 | 0.69 | 1,090 | 1.38 | 1.30 | 352 | 1.52 | 1.07 |
| Annual employment growth | [-0.09, 0.25] | 4,150 | 0.02 | 0.18 | 1,090 | 0.02 | 0.21 | 352 | 0.01 | 0.18 |
| % of firms registered with BRELA | [0, 1] | 4,160 | 0.03 | 0.16 | 1,093 | 0.03 | 0.18 | 353 | 0.05 | 0.22 |
| % of firms with tax ID | [0, 1] | 4,160 | 0.08 | 0.27 | 1,093 | 0.09 | 0.28 | 353 | 0.05 | 0.21 |
| % of firms with business run out of home | [0, 1] | 4,160 | 0.08 | 0.27 | 1,093 | 0.09 | 0.28 | 353 | 0.05 | 0.21 |
| Average monthly value added per firm | 1,000 TZS | 4,160 | 199 | 429 | 1,093 | 252 | 496 | 353 | 256 | 661 |
| Average monthly sales per firm | 1,000 TZS | 4,160 | 383 | 979 | 1,093 | 463 | 1,427 | 353 | 466 | 797 |
| Firm’s age | Year | 4,121 | 6.46 | 6.08 | 1,079 | 6.24 | 5.98 | 348 | 5.43 | 5.35 |
| % of firms with business as full-time | [0, 1] | 4,160 | 0.80 | 0.40 | 1,093 | 0.80 | 0.40 | 353 | 0.83 | 0.38 |
| Keeps accounts in ledger | [0, 1] | 4,160 | 0.39 | 0.49 | 1,093 | 0.51 | 0.50 | 353 | 0.46 | 0.50 |
| Hires paid workers | [0, 1] | 4,160 | 0.07 | 0.26 | 1,093 | 0.16 | 0.37 | 353 | 0.23 | 0.42 |
| >20 customers per day | [0, 1] | 4,160 | 0.28 | 0.45 | 1,093 | 0.29 | 0.46 | 353 | 0.32 | 0.47 |
| Firms powers business with electricity | [0, 1] | 4,160 | 0.09 | 0.28 | 1,093 | 0.35 | 0.48 | 353 | 0.42 | 0.50 |
| Owner savers in formal bank account | [0, 1] | 4,160 | 0.05 | 0.22 | 1,093 | 0.15 | 0.36 | 353 | 0.16 | 0.37 |

**Owner/household characteristics**

| Name of variable | Value unit or range | Observations Mean | S.D. | Observations Mean | S.D. | Observations Mean | S.D. |
|------------------|---------------------|-------------------|-----|-------------------|-----|-------------------|-----|
| Age of owner     | Year                | 4,160             | 36.92 | 10.53 | 1,093 | 37.07 | 11.01 | 353 | 36.97 | 10.62 |
| Whether owner is female | [0, 1] | 4,160 | 0.47 | 0.50 | 1,093 | 0.67 | 0.47 | 353 | 0.68 | 0.47 |
| % firms with business as main source of income | [0, 1] | 4,160 | 0.43 | 0.49 | 1,093 | 0.35 | 0.48 | 353 | 0.26 | 0.44 |
| % firms with farming as main source of income | [0, 1] | 4,160 | 0.24 | 0.43 | 1,093 | 0.06 | 0.25 | 353 | 0.00 | 0.00 |
| % of firms with business as only source of income | [0, 1] | 4,160 | 0.28 | 0.45 | 1,093 | 0.46 | 0.50 | 353 | 0.55 | 0.50 |
| % of firms’ households that are not poor | [0, 1] | 4,160 | 0.41 | 0.49 | 1,093 | 0.53 | 0.50 | 353 | 0.61 | 0.49 |
| % of firms’ households that are moderately poor | [0, 1] | 4,160 | 0.37 | 0.48 | 1,093 | 0.32 | 0.47 | 353 | 0.27 | 0.45 |
| % of firms’ households that are very poor | [0, 1] | 4,160 | 0.22 | 0.41 | 1,093 | 0.15 | 0.36 | 353 | 0.12 | 0.32 |

**Notes:** Full-time employees include business owners and their family members if they work full-time in the firms. BRELA is Tanzania’s Business Registration and Licensing Agency opened in 1999. Household poverty was reported in the survey by an indicator variable equal to zero if the household is not poor, one if the household is moderately poor and two if the household is very poor. The measure of poverty was computed using monthly household income as reported by survey respondents, and the poverty measured in the HBS is considered. TZS denote Tanzanian Shillings and 1000 TZS equivalent to $0.7 US dollars in 2010.

**Source:** Authors’ calculations using the MSME Survey 2010
costs of production. The mean value-added of rural firms reported in Table 3 is about 20 per cent lower than the mean value-added of urban firms in both Dar es Salaam and other urban areas. However, there is significant variation among surveyed firms in monthly value-added in both rural and urban areas, indicated by the high value of the standard deviation (s.d.) in Table 3. We will return to this point in detail later in this paper.

Most MSME firms are young, with a mean age of 6.9 years for rural firms, 6.1 years for urban firms outside Dar es Salaam, and 5.5 years for firms in Dar es Salaam. This is consistent with the findings in Section 1 that most nonfarm jobs created in Tanzania between 2002 and 2012 were created by the informal sector. Table 3 also indicates that 76 per cent of rural businesses operate full time, compared to 82 per cent in urban areas outside Dar es Salaam and 87 per cent in Dar es Salaam. More than 40 per cent of rural business owners report that the business is the owners’ main source of income with a significantly lower share of rural business owners (28%) reporting that the business is the owners’ only source of income. By contrast, 44 per cent and 51 per cent of business owners in other urban and Dar es Salaam report that the enterprise is their only source of income. Only about a quarter of rural business owners report that farming is their main source of income, a fact that supports the finding from the 2011/2012 HBS, that many households (about 10%) in rural areas only participate in the nonfarm economy as their primary employment.

Like their businesses, the owners of these small businesses are also relatively young. For the full sample, the mean age of business owners is roughly 37 years in rural and 36 years in urban areas. In contrast, according to the HBS data, the average age of a rural household’s head is 47 years old and 42 years old for an average urban household’s head.

Finally, the last three rows of Table 3 report the distribution of business owners by the three categories of their household’s income. The measure of these three income categories (very poor, modestly poor and not poor) was computed using monthly household income reported by survey respondents. Poverty appears to be higher among households with MSME owners than among overall rural and urban households. Based on the poverty assessment profile reported by the government (NBS [National Bureau of Statistics], 2013), there are 66.7 per cent of the rural population, 78.3 per cent of the urban population outside of Dar es Salaam and 95.8 per cent of Dar es Salaam’s population lives in nonpoor households. In contrast, in the MSME survey, only 45 per cent of rural MSME owners, 54 per cent of urban MSME owners outside of Dar es Salaam and 62 per cent of MSME

---

**Table 4. Distribution of rural households with and without nonfarm activities (2011/2012)**

|                              | Without nonfarm enterprises | With nonfarm enterprises | Without nonfarm enterprises | With nonfarm enterprises | Nonfarm only |
|------------------------------|-----------------------------|--------------------------|-----------------------------|--------------------------|--------------|
| Share by household numbers   | 61.4                        | 22.1                     | 5.4                         | 5.8                      | 5.4          |
| Share by youth-headed households | 56.9                    | 20.5                     | 4.7                         | 8.7                      | 9.2          |
| Share by population          | 58.4                        | 25.8                     | 6.3                         | 5.1                      | 4.4          |
| Share by employment          | 59.4                        | 26.6                     | 6.6                         | 3.9                      | 3.5          |
| Share by youth employment    | 58.1                        | 24.1                     | 5.3                         | 6.1                      | 6.4          |
| Headcount poverty rate (%)   | 39.3                        | 27.3                     | 31.8                        | 14.2                     | 15.3         |

Notes: The HBS asked individual households whether they have a home business and an ISIC rev4 code is used for assigning sectors to the business. We consider ISIC non-agricultural sectors only as nonfarm enterprises. The employment of non-agriculture is defined by the current primary employment that is not in agriculture. Source: Authors’ calculations based on the data of HBS 2011/2012 (NBS, 2014a).
owners in Dar es Salaam live in nonpoor households. This seems to confirm that in both urban and rural areas, small businesses are often part of a coping strategy for many poor households, while rich households are less likely to choose such small businesses as their main livelihoods.

3.3. Methodologies

The empirical strategy employed in this paper aims to answer to two questions: (1) What determines whether households participate in the nonfarm economy? And (2) what determines whether nonfarm enterprises have the potential to contribute to employment and labour productivity growth? To answer these questions econometric analyses are used.

Descriptive statistics for the HBS and MSME data provide a glimpse of the heterogeneity that is observed across households and firms. The means and standard errors presented in all the descriptive tables of this paper were generated using the sampling design of the two surveys (HBS and MSME).

To address the first question, the rural households in the HBS data are classified into three types based on their family members’ primary employment in the econometric analysis, while more subgroups of households are further classified in the descriptive analysis. The three types of rural households are: (1) farm, in which all family members’ primary employment is agriculture; (2) mixed, indicating that in the same household some family members work in agriculture and others in rural nonfarm economy; and (3) nonfarm, in which all family members work in the rural nonfarm economy as their primary employment. A multinomial probit model is used in the analysis. The choice of a multinomial probit over a multinomial logit arose from the fact that the multinomial probit is better suited to handle correlations and is not bound by the independence of irrelevant alternatives like the multinomial logit. The farm household group is chosen as the comparison category in the regression.

For the MSME dataset, the left-hand side variable is binary which under normal circumstances could be addressed with a simple probit model. However, we must modify our approach due to possible endogeneity issues in some of the right-hand side variables. Dealing with endogeneity in non-linear models (particularly in a probit model) is a straightforward exercise if the endogenous variable is continuous. In such cases, endogeneity is dealt using a control function approach as explained in Wooldridge (2010) and StataCorp (2015). The endogenous variables in our dataset are binary in nature, however. This eliminates the possibility of using a control function approach but raises the possibility of using a bivariate probit model. However, this last approach is also ruled out because it only allows for one endogenous variable while our dataset contains multiple endogenous variables. To address this issue, we have therefore resorted to estimating the probit model using a generalised method of moments (GMM) approach which allows for the endogeneity of multiple variables.

Estimating probit models using GMM is straight-forward if only exogeneous variables are present in the right-hand side, as is shown in StataCorp (2015). However, in the presence of endogeneity the estimation becomes considerably more complicated, as instruments cannot simply be added to the moment conditions of a GMM instrumental variable approach (as one would do for linear models). Doing so for probit or logit models is not possible since neither the conditional expectations nor the linear projections assumed for the linear model apply in the case of probit models. We have therefore followed Wilde (2008) to estimate a two-stage generalised methods of moments (GMM) model which accounts for both the non-linearity of the model and the binary nature of the endogenous variables.

The GMM estimation as proposed by Wilde takes the following form. In the first stage, a reduced form probit model is estimated for each endogenous variable and the residuals are calculated. Having estimated stage 1, a two-step GMM approach is used to estimate the parameters using the correct moment conditions and the necessary adjustments to the standard errors. We refer the reader to Wilde (2008) for details on the specification of the moment conditions. The method proposed by Wilde and as applied in this paper leads to an exactly identified estimation, meaning that the number of instruments equal the number of parameters. Thus, no tests for overidentification of instruments could be conducted for the final estimation. The left-hand side variable for the structural model was defined as follows. It took the value of one if the firm’s value-added per worker (which is used to measure the firm level labour productivity) is greater than the economy-wide labour productivity in the...
trade sector at the national level and zero otherwise. The small firms with high potential are defined as the ‘in-between’ firms following Lewis (1979).

The specifications of the models described above are provided in Sections 3.2 for the HBS survey and Section 4.4 for the MSME survey. Average marginal effects, which can be interpreted as the change in the predicted probability given a one unit change in the right-hand side in the case of continuous variables or a discrete change in the case of categorical variables, are reported. With the exceptions of two variables (firm age and number of employees), all our right-hand side variables are binary in nature. Thus, the marginal effects should be interpreted as discrete changes and, as such, we report average marginal effects. All estimations are done using robust standard errors in accordance to the sampling design. Since the survey was not in any way stratified and subnational units are not representative, we have not clustered the standard errors to any specific subnational location.

4. Characteristics of households with rural nonfarm activities

We begin this section using the 2011/2012 HBS to assess the size of the rural nonfarm economy by the number (or share) of rural households that participate in the nonfarm economy and the differences between households with and without rural nonfarm participation. We then analyse the characteristics of the three categories of households using a multinomial probit model.

4.1. How large is the rural nonfarm economy?

As a starting point, we first classify rural households using the HBS data into with and without nonfarm activities. The classification is based on household members’ primary employment. Like other low-income countries in Africa, most states in Tanzania are predominantly rural (Davis, Di Giuseppe, & Zezza, 2014). Farm activity dominates rural Tanzania – 61.4 per cent of rural households’ members engage only in agricultural activities in 2012 (Table 4), and less than 40 per cent engage in nonfarm activities. Households engaging in rural nonfarm activities can further be classified as farm/nonfarm mixed and nonfarm only households. As shown in Table 4, about 27.5 per cent of total rural households are farm/nonfarm mixed households and about 11 per cent are nonfarm only households.

We also further categorise rural households with nonfarm activities according to whether they have their own nonfarm businesses. This helps us to establish a link between the HBS survey and the MSME survey to be analysed later. According to the HBS, fewer mixed households have their own nonfarm businesses, while almost half of the nonfarm only households own nonfarm businesses. These numbers are comparable to the statistics drawn from the firm data of the MSME survey in Table 3, which shows that 43 per cent of rural nonfarm enterprises are the household’s main income source, but only 28 per cent with businesses report that the business is the only source of income. Finally, we also show in Table 4 that rural households with nonfarm businesses are less likely to be in poverty.

4.2. Characteristics of households in the rural nonfarm economy

Characteristics of the different types of households are explored using a multinomial probit model as discussed in Section 2.3. Equation (1) below describes the specification used in the estimation

\[ y_i = a + \beta_1(H_i) + \beta_2(C_j) + \beta_3(D_r) + \epsilon_i \] (1)

where \( y_i \) is the choice of a given household and takes three values (1 = farm, 2 = mixed, 3 = nonfarm), \( H_i \) is a vector of household characteristics, \( C_j \) is a vector of infrastructure or other community level factors, \( D_r \) is a set of regional dummies. The variables in the vector \( H_i \) include a dummy equal to one if the household is headed by a young person (age 15–34); a dummy equal to one if the household head is female; dummies for the levels of education of the household heads (less than primary as the
comparable variable) and; dummies for farm size defined by cultivated area categorised into four groups: no-land, farms with less than two ha, farms with two to five ha, and farms with more than five ha (no land is the comparison group). Vector $\mathbf{C}_j$ contains a set of variables related to access to infrastructure at the community level and other community level variables including daily public transportation to the regional capital, electricity, mobile phone signal, internet, banks, informal finance, cooperatives, a large employer (for example, a factory), and a weekly market. $\epsilon_i$ is the iid error term.

Table 5 reports the average marginal effects for the three types of households. We begin with the household variables. Being a young household head has a significant and positive effect on being a nonfarm household, with the predicted probability increasing by 3.1 per cent, and has a negative effect at a similar scale on being a mixed household. The level of the household’s head education shows a

| Variable                                      | Farm       | Mixed      | Nonfarm    |
|-----------------------------------------------|------------|------------|------------|
| Youth (age 15–34) as head of a household     | 0.00622    | −0.0370*** | 0.0308***  |
|                                               | (0.0153)   | (0.0127)   | (0.00913)  |
| Female headed household                       | 0.0149     | −0.0217    | 0.00679    |
|                                               | (0.0197)   | (0.0145)   | (0.0109)   |
| (Compared to less than primary) Primary education | −0.0639*** | 0.00566    | 0.0583***  |
|                                               | (0.0156)   | (0.0103)   | (0.0127)   |
| Secondary & higher                           | −0.290***  | 0.103***   | 0.196***   |
|                                               | (0.0259)   | (0.0177)   | (0.0196)   |
| (Compared to no land) Less than 2 ha         | 0.240***   | −0.0283    | −0.221***  |
|                                               | (0.0296)   | (0.0254)   | (0.0162)   |
| 2–5 ha                                        | 0.273***   | −0.0264    | −0.247***  |
|                                               | (0.0318)   | (0.0259)   | (0.0185)   |
| More than 5 ha                               | 0.206***   | 0.0309     | −0.237***  |
|                                               | (0.0377)   | (0.0273)   | (0.0235)   |
| Community variables                           |            |            |            |
| Daily public transport to the regional capital| −0.0298    | −0.00986   | 0.0397**   |
|                                               | (0.0290)   | (0.0179)   | (0.0201)   |
| Electricity access                            | −0.0285    | 0.0258     | 0.00277    |
|                                               | (0.0289)   | (0.0200)   | (0.0171)   |
| Mobile phone signal                           | 0.0678**   | −0.0682*** | 0.000398   |
|                                               | (0.0288)   | (0.0188)   | (0.0187)   |
| Internet access                               | −0.0511    | 0.0365*    | 0.0145     |
|                                               | (0.0337)   | (0.0211)   | (0.0198)   |
| Formal banks                                  | −0.0261    | 0.0236     | 0.00249    |
|                                               | (0.0492)   | (0.0321)   | (0.0401)   |
| Informal financial services                   | 0.0667***  | −0.0282    | −0.0385*** |
|                                               | (0.0248)   | (0.0170)   | (0.0139)   |
| Cooperatives                                  | −0.0253    | 0.0331**   | −0.00781   |
|                                               | (0.0255)   | (0.0145)   | (0.0175)   |
| A major state employer (business or factory)  | −0.0517    | 0.00614    | 0.0455**   |
|                                               | (0.0353)   | (0.0220)   | (0.0223)   |
| Weekly market                                 | −0.0756*** | 0.0409**   | 0.0346**   |
|                                               | (0.0246)   | (0.0144)   | (0.0154)   |
| Observations                                  | 4,053      | 4,053      | 4,053      |

Notes: Standard errors in parentheses. *p < 0.05; **p < 0.01; ***p < 0.001. The average marginal effects (predicted probabilities) based on the multinomial probit regression are reported. In the multinomial probit regression, the farm household group is chosen as the comparison category. See Table 5 in Sosa-Rubi, Galarraga, and Harris (2009) for a similar way to report the result. Source: Authors’ calculation from their estimation results of multinomial probit regression using 2012 Tanzania HBS data.
distinct and contrasting effect on being a farm or a nonfarm household. While both primary and secondary/higher education matter for being a nonfarm household, the more educated the household's head, the larger the effect. Only the higher level of education affects the probability of being a mixed household, as the effect of primary education is insignificant. These results seem to indicate that higher levels of education may be required to obtain nonfarm jobs in rural areas.

Next, we look at farm size. As expected, having a larger farm size decreases the probability of being a nonfarm household by 22–24 per cent. Likewise, a larger farm size is associated with a higher predicted probability of being a farm household by 20.6–27.3 per cent.

Differences among the three types of households are less pronounced for community level variables, perhaps due to the decreased variability that is inherent in these variables. Indeed, public transportation to the regional capital, a proxy for road access, only positively affects the probability of being a nonfarm household. Having a mobile phone signal has the opposite effect on being a farm and a mixed household, positive for the former and negative for the latter, while the effect on being a nonfarm household is insignificant. While the use of electricity for doing business, especially in the manufacturing sector, is important, the variable is not significant for any type of household; this is also true for internet access. This may be because access to electricity or internet at the community level does not necessarily imply access at the household level. Access to informal financial services is associated with a higher probability of being a farm household and a lower probability of being a nonfarm household, suggesting that informal financing is the main channel to borrow money for farm households. The presence of cooperatives is only significant for the effect on being a mixed household with a 3.3 per cent greater predicted probability. As expected, the presence of a large employer is associated with a higher predicted probability of being a nonfarm household by 4.5 per cent, but does not influence being a mixed household. Finally, having access to weekly markets is the only variable that is significant across all types of households. Access to markets reduces the predicted probability of being a farm household by 7.5 per cent and increases the predicted probability of being mixed and nonfarm households by 4 per cent and 3.4 per cent respectively. The mixed results regarding the role of infrastructure are puzzling. As noted, this may be because the community level variables are too ‘rough’ a proxy for access at the household level. However, the lack of significance of these variables may also be associated with the small scale of rural nonfarm enterprises. According to Tybout (2000), low levels of economic density and interaction may lead to small, diffuse pockets of demand, which in turn result in small, localised production and services. We revisit this issue in the next section using the MSME data.

5. Characteristics of rural nonfarm enterprises and their owners – an analysis at the firm level using MSME survey data

A vibrant rural nonfarm sector can play an important role in rural transformation. To understand the extent to which the rural nonfarm sector can play a role in labour productivity growth and poverty reduction in rural areas, the MSME survey data is used to examine the motivations of business owners in the rural nonfarm sector as well as the characteristics of their businesses. In a previous paper, Diao, Kweka, McMillan and Qureshi (2017) identify a group of MSMEs that can be considered members of what Arthur Lewis (1979) referred to as the in-between sector. According to Lewis (1979), these firms play an important role in the transformation process. Lewis (1979) uses the term in-between to signal that these firms are not just petty traders, rather, they often look more like formal firms and provide important goods and services. Diao, Kweka, et al. (2017) show that rural enterprises are on average slightly less productive than their urban counter-parts (this is confirmed in Table 3 of Section 2 in this paper) but they do not explore in detail the characteristics of rural enterprises or rural entrepreneurs.

This section begins with a description of the location and industrial composition of MSMEs. It follows with an exploration of the extent to which rural entrepreneurs appear to be subsistence or growth-oriented. To analyse this issue, we use the following data from the MSME survey: (i) self-reported motivations for business ownership; (ii) the productive heterogeneity of MSMEs and; (iii) employment growth in MSMEs.
5.1. Industrial and geographic composition of MSMEs

Table 6 reports the distribution of employment and the number of MSMEs by rural, other urban and Dar es Salaam, compared with the distribution of population in the three locations. While more than 67 per cent of the population lives in rural areas, rural MSMEs account for 52 per cent of total MSME employment. In urban areas, the distribution of MSME employment/firms and distribution of population seem to be similar in Dar es Salaam and other urban areas. 15.8 per cent of MSME employment and 17.3 per cent of MSME firms are in Dar es Salaam, where 12.2 per cent of the national population resides. Likewise, 32.6 per cent of MSME employment and 30.7 per cent of MSME firms are in other urban areas, which contain 20.4 per cent of the population (Table 6).

Table 7 reports the industrial distribution of MSMEs by rural, other urban and Dar es Salaam. Although the MSMEs operate in a wide range of activities, the bulk of these activities can be classified as trade services (80%) and manufacturing (15%). However, more rural firms (19.8%) engage in manufacturing than urban firms (10.1% in other urban and 7.2% in Dar es Salaam). Seventy-two per cent of manufacturing MSMEs are in rural areas while 52 per cent of trade service MSMEs are in rural areas. This is an expected pattern, as small manufacturing firms operate mainly in food processing, which has strong links to agriculture. Without further information, however, it is not possible to identify exactly what these linkages are and how they work. This is an important area for future research. More firms are in the trade services in Dar es Salaam (87.6%) than in other urban (83.0%), which is clearly driven by demand for tradable goods.

5.2. Self-reported motivations of small business owners

The MSME survey includes three questions designed to elicit the reasons for opening a business. Responses to such self-reported motivations for a business could help us assess the extent to which rural entrepreneurs are in business solely for the purposes of survival or aiming to grow. The responses to these questions are tabulated using sample weights in Tables 8–10.

Table 6. Distribution of population and MSMEs (weighted, percentage)

|                  | Population | MSME employment | Number of business |
|------------------|------------|-----------------|--------------------|
|                  | All        | All MSMEs       |                    |
| Rural            | 67.4       | 51.6            | 52.1               |
| Other urban      | 20.4       | 32.6            | 30.7               |
| Dar Es Salaam    | 12.2       | 15.8            | 17.3               |
| Total            | 100        | 100             | 100                |

Source: Population is from HBS (2012) and MSME employment and number are from MSME survey (2010).

Table 7. Sectoral distribution of rural and urban MSME firms in the survey (weighted, percentage)

| Percentage of total in each location | Percentage in each sector (National total by sector = 100) |
|--------------------------------------|----------------------------------------------------------|
|                                      | Rural          | Other urban | Dar es Salaam | National       | Rural          | Other urban | Dar es Salaam |
|                                      | Manufacturing | Trade services | Others | Manufacturing | Other urban | Trade services | Others |
| Total                                | 100           | 100         | 100        | 100           | 54.2         | 31.2        | 14.6          |
| Manufacturing                       | 19.8          | 10.1        | 7.2        | 14.9          | 71.9         | 21.1        | 7.1           |
| Trade services                      | 76.5          | 83.0        | 87.6       | 80.2          | 51.7         | 32.3        | 16.0          |
| Others                              | 3.8           | 6.9         | 5.2        | 4.9           | 41.1         | 43.5        | 15.4          |

Source: Authors’ calculations using the MSME Survey 2010.
The first question is: ‘What was your main occupation before you started this business?’ As shown in Table 8, the biggest difference between rural and urban entrepreneurs is that 56.5 per cent of rural entrepreneurs report that their main occupation prior to starting the business was farming compared to 19.3 per cent in urban areas outside Dar es Salaam. Very few respondents (4.8%) in rural areas report that they were unemployed prior to starting the business; this is not true in urban areas where 11.3 per cent and 9.7 per cent of MSME owners in other urban and Dar es Salaam report that they were unemployed before starting their business. Unlike in rural areas, urban business owners are much more likely to report that they were previously employed in a private company or running a similar sized business in another line of business. It is also much more common for urban business owners to report that they were previously a housewife or homemaker (26.6% in other urban and 34.1% in Dar es Salaam) than for rural respondents (12.3%).

The second question is: ‘For what reasons did you choose your line of business?’ In Table 9, the firms responding to this question are grouped into three broad sectors: manufacturing, trade services and other services, by rural, other urban and Dar es Salaam. In rural areas, half of all business owners say that the reason they chose their line of business is because they saw a market opportunity. This response is similar for firms in manufacturing and trade services. However, this response is less common in Dar es Salaam and other urban areas. The second most common reason for operating in a line of business in rural areas is that the owners’ capital could only finance that line of business; this response is more common in urban than rural areas possibly indicating that capital constraints are more severe in urban areas. The third most common reason for choosing a line of business in rural areas was prior experience in that line of business, although shares for this reason are much lower than the two previous reasons.

The third question is: ‘If you were offered a full-time salary paying job, would you take it?’ Responses to this question are reported in Table 10 and indicate that only 46.6 per cent of all small business owners would leave their current business for a full time salaried position, but the share is higher in rural areas (47.8%) and other urban areas (48.6%) than in Dar es Salaam (37.7%). Approximately 64 per cent of all respondents who would prefer a full time salaried job say they

Table 8. Occupation prior to starting business of MSMEs (weighted, percentage)

| Occupation Prior to Starting Business | Rural | Other Urban | Dar es Salaam | Total |
|--------------------------------------|-------|-------------|---------------|-------|
| Unemployed                           | 4.8   | 11.3        | 9.7           | 7.6   |
| Housewife (home maker)               | 12.3  | 26.6        | 34.1          | 20.0  |
| In education, at various levels      | 3.2   | 5.5         | 5.4           | 4.2   |
| Employed in large private enterprise in similar business | 0.4 | 1.7 | 2.9 | 1.2 |
| Employed in large private enterprise in a different business | 1.7 | 4.7 | 6.9 | 3.4 |
| Employed in a similar sized private business in the same line of business | 0.6 | 1.5 | 1.9 | 1.0 |
| Employed in a similar sized private business in another line of business | 0.6 | 0.8 | 3.0 | 1.0 |
| Ran a similar sized business in the same line of business | 0.9 | 2.0 | 1.9 | 1.4 |
| Ran a similar sized enterprise in another line of business | 9.5 | 16.4 | 20.0 | 13.2 |
| Civil servant/employed by the government | 1.8 | 2.1 | 6.0 | 2.5 |
| I was employed by some individual   | 0.6   | 2.2         | 1.9           | 1.3   |
| Rearing of cattle                   | 0.5   | 0.4         | 0.2           | 0.4   |
| Farming                             | 56.5  | 19.3        | 0.7           | 36.7  |
| I was selling food                  | 0.6   | 1.4         | 1.1           | 0.9   |
| Others                              | 4.8   | 3.5         | 2.4           | 4.0   |
| None                                | 1.2   | 0.6         | 2.0           | 1.1   |

Notes: This table is prepared based on the question ‘what was your main occupation before you started this business?’ in the MSME survey, and a unique answer is provided by individual MSME owners. The sum of each column in the table is 100.

Source: Authors’ calculations using the MSME Survey 2010.
| Reasons for Business Choice                                      | Rural Enterprise | Other Urban Enterprises | Enterprise in Dar es Salaam |
|---------------------------------------------------------------|------------------|------------------------|-----------------------------|
|                                                               | Manufacturing    | Trade services         | Rural total                 |
| All MSMEs                                                     | 25.0             | 15.2                   | 18.3                        |
| I had previous experience in this line                       | 20.6             | 13.4                   | 14.8                        |
| Friends/relatives are in this line                           | 48.2             | 51.6                   | 50.0                        |
| My capital could only finance this business                  | 36.1             | 42.1                   | 41.8                        |
| No apparent reason                                           | 2.9              | 6.0                    | 4.5                         |
| I could start business gradually                             | 0.0              | 0.1                    | 0.1                         |
| Goods are easy to manufacture and sell                       | 1.4              | 2.0                    | 2.0                         |
| I just wanted to be near my house                            | 0.8              | 1.0                    | 0.9                         |
| I have been trained in it, I am an expert                    | 1.6              | 0.4                    | 0.6                         |
| Goods are available                                          | 0.3              | 0.4                    | 0.5                         |
| I perceived it to be profitable                               | 1.3              | 1.6                    | 1.7                         |
| I liked it                                                    | 0.7              | 1.0                    | 1.3                         |
| Business does not have many problems                         | 1.4              | 0.6                    | 0.7                         |
| Other                                                         | 1.0              | 2.5                    | 2.2                         |
| None                                                          | 0.3              | 1.3                    | 0.8                         |

Notes: Multiple answers are allowed for individual MSME owners.
Source: Authors’ calculations using the MSME Survey 2010.
would like to work for the government, with 68.6 per cent and 62.9 per cent in rural areas and other urban areas respectively but only 44.6 per cent in Dar es Salaam, where more government jobs are concentrated. The responses from rural and other urban MSME owners are consistent with results reported in Banerjee and Duflo’s analysis of the economic lives of the poor (Banerjee & Duflo, 2007).

Large private companies are more attractive to small business owners in Dar es Salaam than in other places. The predominant reason for preferring a full time salaried position is better security of income.

5.3. The productive heterogeneity of rural enterprises

The kernel densities of the log of value added per worker, which is defined as firms’ labour productivity, is used to examine the productive heterogeneity of MSMEs. Value added is computed as the firm’s average monthly sales minus the firms’ average monthly costs of production, and seasonality is taken into consideration in the calculation. Only full-time employees (including owners of the firms) are considered in calculating value-added per worker or labour productivity for individual firms. The kernel densities of labour productivity reveal two important features of the MSME firms.

First, there is a significant degree of productive heterogeneity among both rural, urban and Dar es Salaam enterprises. This can be seen by examining the density of the log of value added per worker in Figure 1. Surprisingly, the distribution of the log of value added per worker or labour productivity for rural firms is almost identical to the distribution for urban firms. In fact, stochastic dominance test rejects the hypothesis that the rural and urban distributions are not identical.8 One reason for this may be the fact that medium sized enterprises that are mainly in urban areas appear to be under-sampled in the MSME survey discussed in Section 2.

In Figure 1 the vertical lines represent average labour productivity in Tanzania’s economy in 2010 in the agricultural sector (the far-left line), the trade services sector (the middle-line) and the manufacturing sector (the far-right line). Economy-wide labour productivity is calculated using national accounts data and census data; since 1997 national accounts data make every attempt to include the informal

| Table 10. Job satisfaction in MSME survey (weighted, percentage) |
|---------------------------------------------------------------|
| All MSMEs | Rural | Other urban | Dar es Salaam | Total |
| --- | --- | --- | --- | --- |
| If you were offered a full-time salary paying job, would you take it? | 47.8 | 48.6 | 37.7 | 46.6 |
| Who would you rather work for? | | | | |
| A large private company | 17.9 | 27.4 | 43.1 | 24.0 |
| Government | 68.6 | 62.9 | 44.6 | 63.9 |
| Someone else’s business | 10.5 | 6.2 | 10.8 | 9.1 |
| Anywhere | 3.0 | 3.4 | 1.5 | 3.0 |
| And why do you say that? | | | | |
| Better security of income | 81.7 | 83.8 | 81.4 | 82.3 |
| Shorter hours | 5.1 | 5.7 | 3.5 | 5.1 |
| Less risk | 1.8 | 1.8 | 3.0 | 1.9 |
| To get pension | 1.5 | 1.6 | 1.4 | 1.0 |
| I am less educated | 2.2 | 0.9 | 1.9 | 1.8 |
| They listen to the opinions of the employees | 1.1 | 1.0 | 1.0 | 1.0 |
| As long as I get a living | 0.6 | 0.4 | 0.5 | 0.5 |
| Job security | 2.1 | 0.6 | 1.4 | 1.4 |
| Others | 2.2 | 2.3 | 9.9 | 3.1 |
| None/Nothing | 1.7 | 2.1 | 0.2 | 1.6 |

Notes: This table is prepared based on three questions: (1) ‘If you were offered a full-time salary paying job, would you take it?’ (2) ‘Who would you rather work for?’ and (3) ‘Why do you say that?’ A unique answer is provided by individual MSME owners to each of the last two questions. Rural, urban and national total MSMEs, MSMEs owned by youth and MSMEs owned by other adults for the sum of these two questions are 100 respectively. Source: Authors’ calculations using the MSME Survey 2010.
However, in practice it is difficult to accurately measure informal sector activity and so it is likely that economy-wide estimates of labour productivity are biased toward the formal sector.

Figure 1 reveals that a little over half of the firms in the MSME sector have labour productivity levels higher than the average labour productivity in the agricultural sector and this is true in all three locations. This is not surprising and is consistent with evidence presented by Diao, Kweka, et al. (2017), who show that labour productivity for many MSMEs is consistently higher than average labour productivity in the agricultural sector. It is also true that around 25 per cent of rural and urban MSMEs have labour productivity higher than average labour productivity in the services sector – a sector most MSMEs belong to. In fact, as shown in Diao, Kweka, et al. (2017), this group of 25 per cent of small firms accounts for 77 per cent of total value-added produced by the whole MSME sector. In other words, the remaining 75 per cent of MSMEs account for less than 25 per cent of the value added generated by the MSME sector. These results underscore the productive heterogeneity of MSMEs in both rural and urban areas. They also raise the possibility of a growth strategy focused on these most productive firms. This is not to say that the remaining firms should not be part of a strategy for alleviating poverty, perhaps they should. Our point is that the productive heterogeneity most likely calls for different strategies for different types of firms.

6. Using the MSME survey to identify ‘high potential’ rural enterprises

If we accept that some rural MSMEs have more potential to contribute to rural transformation in Tanzania than others, we are left with the question of how to identify those with potential. This is a complicated problem not least because a properly designed mechanism should be immune from manipulation. We do not pretend to solve it. Instead, what follows is meant to be illustrative of the way in which we are thinking about the problem. We use a productivity cutoff to distinguish in-between MSMEs from the rest of the MSMEs and then we look for readily observable characteristics of these highly productive MSMEs that might be used for targeting. In practice, it would be important to use readily observable characteristics that cannot be manipulated or are too costly to be manipulated by firms.

For the purposes of this exercise, we define the in-between firms as those with labour productivity greater than economy-wide labour productivity in trade services. Using this criterion, we identify 1334 rural firms in the MSME sample that can be classified as belonging to the in-between sector. Having
selected our in-between firms, we use the GMM probit analysis previously described to identify characteristics of in-between firms. We use a host of business and owner characteristics that have been used and tested in the literature.

Prior to discussing the results of the GMM probit regressions, we must first identify the endogenous variables and present the instruments used for the reduced form probit estimations. Three endogenous variables were identified: the owner views the business as growing, the firm has regional customers, and the number of daily customers is more than 20. The first two variables were instrumented using the following variables: whether the business was located at home, whether the firm advertised, whether the firm had a business plan, and whether the firm regularly sends or receives money. The model was run using a two-stage GMM with robust standard errors. Results of the first-stage estimations are available upon request.

Table 11 presents the GMM probit results of the increases in predicted probability (the average marginal effect) of being an in-between sector firm in rural and urban areas. Results for the owners’ personal characteristics suggest that a female-headed business is associated with a decrease in the probability of being in-between in both rural and urban areas; probabilities decrease by 6.4 and 9.5 per cent, respectively. Owners that perceive their businesses to be growing observe gains in the predicted probability of being in-between by 4.8 per cent and 6.4 per cent nationally and in rural areas, respectively. These results are intuitive; business owners who are optimistic about their firm’s future and potential are more likely to be driven to achieve success and to use resources productively. Being a member of a business association increases the probability of being an in-between firm nationally and in urban areas, but not in rural areas, possibly due to the low participation rate of rural firms in such associations. Education was not found to be significant anywhere, likely due to lack of variation in education among business owners.

The second panel of Table 11 shows varying levels of significance between locations, which is expected. A one-year increase in a firm’s age has a small effect on the probability of being an in-between firm in the country as a whole and in rural areas, but not in urban areas. A one-unit increase in the number of employees reduces the probability of being an in-between firm by around 5 per cent consistently in both rural and urban areas. On the other hand, operating full time is associated with an increase in the probability of being in-between by 6.2 per cent nationally and 6.7 per cent in rural areas, but is not significant in urban areas possibly for two reasons. First, the survey is designed to capture small businesses, of which many in urban areas may be part-time. Second, unless the businesses are large enough, running a full-time small business in an urban area probably carries a higher opportunity cost if it means not finding a job. Keeping written accounts is significant with an increase in the predicted probability of being in between of around 6.5 per cent in all locations. There are increases in the predicted probability of being in-between for firms which have licenses, with a larger marginal effect in urban areas (6%) than in rural areas (3.4%). The increases in the predicted probability of being in the in-between sector from having regional customers is significant only in rural areas. However, firms that have a daily number of customers greater than 20 are between 7.4 per cent (rural areas) and 6.2 per cent (urban areas) more likely to be in the in-between sector.

The variables associated with the external conditions of doing business, infrastructure and technology, are presented in the third panel of Table 11. Using a mobile phone increases the predicted probability of being in-between by 4.3 per cent in rural areas and 6.2 per cent in urban areas. Whether the business uses electricity to light their businesses is important in rural areas, increasing the predicted probability by 5.6 per cent, but not in urban areas, possibly again due to the lack of variability in electricity access in urban areas. We also include three financing variables in the regression (the last panel of Table 11), and all three variables are related to the ways business owners allocated their profits. It turns out all three variables are insignificant, possibly due to lack of variability in these variables, or they fail to accurately capture firms’ investment behaviour.

7. Summary and policy implications

Although Tanzania remains heavily rural, the composition of economic activity in rural areas has changed significantly over the past decade and a half. Between 2002 and 2012, the share of the rural
Table 11. GMM probit results for probability of being in-between rural and urban enterprises – average marginal effects

| Variable                                | National | Urban | Rural          |
|-----------------------------------------|----------|-------|----------------|
| **Owner’s characteristics**             |          |       |                |
| Education (completed secondary or higher) | 0.00852  | −0.00511 | 0.0261       |
|                                         | (0.0220) | (0.0302) | (0.0307)      |
| Female                                  | −0.0740*** | −0.0954*** | −0.0649***  |
|                                         | (0.0148) | (0.0276) | (0.0161)      |
| Owner would leave for salaried job       | −0.0330** | −0.0484** | −0.0206      |
|                                         | (0.0139) | (0.0241) | (0.0145)      |
| Owner is a Member of a Business Association | 0.0777*** | 0.120*** | 0.00183      |
|                                         | (0.0280) | (0.0382) | (0.0340)      |
| Saw Business as a Market Opportunity     | 0.0198   | 0.0374 | 0.00821       |
|                                         | (0.0137) | (0.0239) | (0.0144)      |
| Views business as growing                | 0.0485*** | 0.0326  | 0.0645***     |
|                                         | (0.0149) | (0.0270) | (0.0154)      |
| **Business characteristics**             |          |       |                |
| Firm’s age                              | 0.00244** | 0.00132 | 0.00309***    |
|                                         | (0.00100) | (0.00185) | (0.00106)    |
| Firm Size (number of employees)          | −0.0508*** | −0.0550*** | −0.0494***  |
|                                         | (0.00990) | (0.0124) | (0.0172)      |
| Business runs as Full time               | 0.0622*** | 0.0438  | 0.0673***     |
|                                         | (0.0187) | (0.0332) | (0.0211)      |
| Firm keeps Written accounts              | 0.0643*** | 0.0627** | 0.0678***     |
|                                         | (0.0154) | (0.0271) | (0.0160)      |
| Firm has a License                       | 0.0467*** | 0.0600** | 0.0337*       |
|                                         | (0.0169) | (0.0287) | (0.0197)      |
| Firm has Regional Customers              | 0.0188   | −0.00993 | 0.0421***    |
|                                         | (0.0176) | (0.0285) | (0.0197)      |
| Number of daily customers is more than 20 | 0.0686*** | 0.0622*** | 0.0742***    |
|                                         | (0.0138) | (0.0241) | (0.0146)      |
| Firm’s suppliers are small traders        | 0.0151   | 0.0269  | 0.00865       |
|                                         | (0.0139) | (0.0245) | (0.0150)      |
| Firm’s Suppliers are Nationwide          | 0.0223   | 0.0523  | −0.0514      |
|                                         | (0.0302) | (0.0390) | (0.0496)      |
| **Infrastructure and technology**        |          |       |                |
| Owner using Mobile to Conduct Business    | 0.0529*** | 0.0617** | 0.0427***     |
|                                         | (0.0157) | (0.0282) | (0.0156)      |
| Fir Owner has a Calculator               | 0.0235   | 0.0370  | 0.00891       |
|                                         | (0.0182) | (0.0312) | (0.0201)      |
| Business uses Electricity to Light Business | 0.0493*** | 0.0433  | 0.0560**      |
|                                         | (0.0187) | (0.0275) | (0.0233)      |
| **Financing**                            |          |       |                |
| Owner uses profit to Expand Business      | 0.0204   | 0.0287  | 0.0194        |
|                                         | (0.0178) | (0.0329) | (0.0173)      |
| Owner uses profits to Buy Stocks in Advance | 0.00483  | −0.0100 | 0.0150        |
|                                         | (0.0145) | (0.0256) | (0.0145)      |
| Owners uses Profits to Invest in Buildings and Land | 0.0291  | 0.0148  | 0.0360        |
|                                         | (0.0262) | (0.0498) | (0.0289)      |
| Observations                            | 5,548    | 1,427  | 4,121         |

Notes: Standard errors in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1. Dependent variable is a binary variable which takes the value of one if the firm is in-between and zero otherwise. Firms in the ‘in-between’ category satisfy the following conditions: labour productivity is higher than economy-wide labour productivity in trade. Source: Authors’ estimation using MSME data.
labour force working in nonfarm employment tripled going from 6.8 per cent to 20.5 per cent. Moreover, in 2011/2012, more than one-third of rural households participated in the rural nonfarm economy and 11.2 per cent of rural households reported that working members of the household had primary employment only in the nonfarm economy.

The heads of ‘nonfarm only’ rural households tend to be younger and more educated, while the heads’ gender does not appear to influence the likelihood of being a nonfarm household. Education of the household head is also a determinant of the likelihood that a household participates in the nonfarm sector; a primary education increases the probability of engaging in nonfarm activities by 5.8 per cent and a secondary education increases the likelihood of engaging in the nonfarm sector by 16.9 per cent. Among a set of selected community level variables, households in communities with access to daily public transportation or a weekly market are more likely to participate in rural nonfarm activities. Consistent with these results, we find that rural households with nonfarm activities are less likely to be poor. However, it is still true that around 15 per cent of rural households whose primary source of income is the nonfarm economy have incomes that place them below the poverty line. The implication is that some nonfarm activities must be very unproductive. By extension, although these activities help families to survive, it would be unrealistic to expect them to contribute significantly to rural transformation.

To explore the nature of the nonfarm businesses owned by rural households in Tanzania, we use Tanzania’s first nationally representative survey of micro, small and medium sized enterprises. Roughly 20 per cent of these businesses operate in the manufacturing sector – more than double the share in urban areas – the rest of the businesses operate in the services sector. Labour productivity among these businesses is extremely heterogeneous with roughly half having labour productivity lower than average labour productivity in agriculture. Using a probit specification we show that operating full time, keeping written accounts and using electricity to run the business are all positively correlated with labour productivity.

We conclude that policies designed to stimulate rural transformation must take into account the heterogeneity of the rural nonfarm sector. Unless this heterogeneity is understood, policies designed to stimulate rural transformation are likely to disappoint. Of course, rural nonfarm activities help to generate income and reduce the risks associated with agricultural production for many rural households. These activities should be supported as part of a poverty reduction strategy. But we should not expect the large majority of these activities to transform rural livelihoods. For this to happen, it will be important to target the firms with the potential for employment and labour productivity growth.

Acknowledgements
This work was undertaken as part of the CGIAR Research Program on Policies, Institutions, and Markets (PIM) led by IFPRI. Funding for the study was provided by PIM. We are grateful to two anonymous reviewers for useful comments on previous versions of the paper, and to Peixun Fang and Jed Silver for excellent research assistance. Finally, Xinshen and Margaret would like to dedicate this article to the memory of Eduardo Magalhaes who sadly and prematurely passed away in August 2017.

Funding
This paper has benefited from funding from the CGIAR Research Program on Policies, Institutions, and Markets (PIM) led by IFPRI.

Disclosure statement
No potential conflict of interest was reported by the authors.
Notes

1. Although other nationally representative surveys of MSMEs exist, the data collected in these surveys is minimal and do not allow for the calculation of labour productivity.
2. For example, the nationally representative GEMINI surveys conducted in the early 1990s and funded by the United States Agency for International Development in a handful of African countries only collected location and employment data for the nationally representative sample. For a much smaller group of around 250 firms per country, more detailed information was collected but still not enough information to calculate value added and thus labour productivity.
3. For example, the World Bank Enterprise Survey conducted in Tanzania in January 2013–August 2014 interviewed 813 firms’ owners and top managers including 514 small firms.
4. The survey used in Jin and Deininger (2009) covers only rural enterprises (1239) and rural households (1610) in Tanzania and cannot be used for a comparison to urban enterprises.
5. BREA is Tanzania’s Business Registrations and Licensing Agency. It is a Government Executive Agency and was established on the 28 October 1999. The aim of the agency is to ensure that businesses operate in accordance with regulations and to ensure that businesses follow ‘sound principles’.
6. See the definition in the notes of Table 3.
7. Stata estimates these models using the ivprobit.
8. A stochastic dominance test compares the cumulative density function of the log of the value added in rural and urban areas.
   To establish whether the two curves originate from the same distribution, we have used the Kolmogorov-Smirnov test of equality of distribution. We were not able to reject the null hypothesis that the distributions of urban and rural value added were the same (p-value = 0.9).

References

Banerjee, A. V., Breza, E., Duflo, E., & Kinnan, C. (2015). Do Credit Constraints Limit Entrepreneurship? Heterogeneity in the Returns to Microfinance. NBER Conference Paper. Cambridge, MA: National Bureau of Economic Research. http://conference.nber.org/confer/2015/SI2015/PRPD/Kinnan.pdf.
Banerjee, A. V., & Duflo, E. (2007). The economic lives of the poor. The Journal of Economic Perspectives: A Journal of the American Economic Association, 21(1), 141–168.
Beck, T. H. L. (2007). Financing constraints of SMEs in developing countries: Evidence, determinants and solutions. In Financing innovation-oriented businesses to promote entrepreneurship. Unknown Publisher.
Bezu, S., & Barrett, C. (2012). Employment dynamics in the rural nonfarm sector in Ethiopia: Do the poor have time on their side? Journal of Development Studies, 48(9), 1223–1240.
Davis, B., Di Giuseppe, S., & Zeeza, A. (2014). Income diversification patterns in rural Sub-Saharan Africa: Reassessing the evidence (Policy Research Working Paper 7108). Washington, DC: World Bank.
De Brauw, A., Mueller, V., & Lee, H. L. (2014). The role of rural–Urban migration in the structural transformation of Sub-Saharan Africa. World Development, 63(2014), 33–42.
Diao, X., Kweka, J., McMillan, M., & Qureshi, Z. (2017). Economic transformation from the bottom up: Evidence from Tanzania (IFPRI Discussion Paper 1603). Washington, DC: International Food Policy Research Institute.
Diao, X., McMillan, M., & Rodrik, D. (2017). The recent growth boom in developing economies: A structural-change perspective (NBER Working Paper 23132). Cambridge, MA: National Bureau of Economic Research.
Fafchamps, M., McKenzie, D., Quinn, S., & Woodruff, C. (2014). When is capital enough to get female enterprises growing? Evidence from a randomized experiment in Ghana. Journal of Development Economics, 106(January), 211–226.
Gibson, J., & Olivia, S. (2010). The effect of infrastructure access and quality on nonfarm enterprises in rural Indonesia. World Development, 38(5), 717–726.
Gollin, D., & Rogerson, R. (2009). The greatest of all improvements: Roads, agriculture, and economic development in Africa. Williamstown, MA: Department of Economics, a mimeo, Williams College.
Grimm, M., Krüger, J., & Lay, J. (2011). Barriers to entry and returns to capital in informal activities: Evidence from Sub-Saharan Africa. Review of Income and Wealth, 57(1), 27–53.
Haggblade, S., Hazell, P. B. R., & Brown, J. (1989). Farm-nonfarm linkages in rural sub-saharan africa. World Development, 17 (August), 1173–1202.
Haggblade, S., Hazell, P., & Reardon, T. (2010). The rural nonfarm economy: Prospects for growth and poverty reduction. World Development, 38(1), 1429–1441.
Jin, S., & Deininger, K. (2009). Key constraints for rural nonfarm activity in Tanzania: Combining investment climate and household surveys. Journal of African Economies, 18(2), 319–361.
La Porta, R., & Shleifer, A. (2011). The unofficial economy in Africa (NBER Working Paper 16821). Cambridge, MA: National Bureau of Economic Research.
La Porta, R., & Shleifer, A. (2014). Informality and development (NBER Working Paper 20205). Cambridge, MA: National Bureau of Economic Research.
Lewis, W. A. (1979). The dual economy revisited. The Manchester School, 47(3), 211–229.
Li, Y., & Rama, M. (2015). Firm dynamics, productivity growth, and job creation in developing countries: The role of micro- and small enterprises. The World Bank Research Observer, 30(1), 3–38.

McKenzie, D., & Woodruff, C. (2008). Experimental evidence on returns to capital and access to finance in Mexico. World Bank Economic Review, 22(3), 457–482.

McKenzie, D. J. (2015). Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition (World Bank Policy Research Working Paper 7391). Washington, DC: World Bank Group.

McPherson, M. (1996). Growth of micro and small enterprises in Southern Africa. Journal of Development Economics, 48, 253–277.

Mead, D. C. (1994). The contribution of small enterprises to employment growth in Southern and Eastern Africa. World Development, 22(12), 1881–1894.

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

Nagler, P., & Naudé, W. (2017). Nonfarm enterprises in rural Africa: New empirical evidence. Food Policy, 67(C), 175–191.

National Bureau of Statistics. (2006). Tanzania 2002 census: Analytic report, Volume X. Dar es Salaam: National Bureau of Statistics, Ministry of Planning, Economy, and Empowerment, Tanzania.

National Bureau of Statistics. (2007). Analytic Report for Employment and Earnings Survey 2002. Dar es Salaam: National Bureau of Statistics, Ministry of Planning, Economy, and Empowerment, Tanzania.

National Bureau of Statistics. (2011a). National household budget survey 2000-2001. Dar es Salaam: National Bureau of Statistics, Ministry of Finance, Tanzania.

National Bureau of Statistics. (2011b). Integrated labor force survey 2014, analytical report. Dar es Salaam: National Bureau of Statistics, Ministry of Finance, Tanzania.

National Bureau of Statistics. (2013). Key findings of 2011/12 household budget survey Tanzania Mainland. Dar es Salaam: National Bureau of Statistics, Ministry of Finance, Tanzania.

National Bureau of Statistics. (2014a). Household budget survey: Main report, 2011/12. Dar es Salaam: National Bureau of Statistics, Ministry of Finance, Tanzania.

National Bureau of Statistics. (2014b). Tanzania 2012 census: basic demographic and socio-economic profile. Dar es Salaam and Zanzibar: National Bureau of Statistics, Ministry of Finance, Office of Chief Government Statistician, Ministry of State, President’s Office, State House and Good Governance, Tanzania.

National Bureau of Statistics. (2014c). National accounts of Tanzania Mainland 2001–2013. Dar es Salaam: National Bureau of Statistics, Ministry of Finance, Tanzania.

National Bureau of Statistics. (2014d). Revised national accounts estimates for Tanzania Mainland, Base Year, 2007. Dar es Salaam: National Bureau of Statistics, Ministry of Finance, Tanzania.

National Bureau of Statistics. (2014e). Formal Sector Employment and Earnings Survey: Analytical Report 2013. Dar es Salaam National Bureau of Statistics, Ministry of Finance, Tanzania.

National Bureau of Statistics. (2015). Integrated labor force survey 2014, analytical report. Dar es Salaam: National Bureau of Statistics, Ministry of Finance, Tanzania.

Nordman, C., Rakotomana, F., & Roubaud, F. (2016). Informal versus formal: A panel data analysis of earnings gaps in Madagascar. World Development, 86, 1–17.

Raj, R., & Sen, K. (2013). How important are credit constraints for small firm growth? Evidence from the Indian informal manufacturing sector (IGC Working Paper). London: International Growth Center.

Rijkers, B., & Costa, R. (2012). Gender and rural nonfarm entrepreneurship. World Development, 40(12), 2411–2426.

Rijkers, B., Soderhom, M., & Loening, J. L. (2010). A rural–Urban comparison of manufacturing enterprise performance in Ethiopia. World Development, 38(9), 1278–1296.

Sosa-Rubi, S. G., Galárraga, O., & Harris, J. E. (2009). Heterogeneous impact of the “Seguro Popular” Program on the utilization of obstetrical services in Mexico, 2001–2006: A multinomial probit model with a discrete endogenous variable. Journal of Health Economics, 28(1), 20–34.

StataCorp. (2015). Stata 14 base reference manual. College Station, TX: Stata Press.

Tybout, J. R. (2000). Manufacturing firms in developing countries: How well do they do, and why? Journal of Economic Literature, 38(1), 11–44.

Wilde, J. (2008). A note on GMM estimation of probit models with endogenous regressors. Statistical Papers, 49, 471–484.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data (2nd ed.). Cambridge, MA: MIT Press.
Appendix

### Table A1. Summary statistics of main variables of 2012 HBS used in the regression – rural

| Variable                                      | Number of HHs or EAs | Means  | SE   | LL   | UL   |
|-----------------------------------------------|----------------------|--------|------|------|------|
| Youths headed (between ages 15–34)            | 4,130                | 0.268  | 0.010| 0.249| 0.287|
| Female headed                                 | 4,130                | 0.241  | 0.009| 0.222| 0.259|
| No Primary Education                          | 4,130                | 0.270  | 0.013| 0.243| 0.296|
| Completed Primary Education                   | 4,130                | 0.673  | 0.012| 0.649| 0.697|
| Completed Secondary Education or more         | 4,130                | 0.058  | 0.008| 0.043| 0.073|
| Number of Youth (15–34)                       | 4,130                | 1.577  | 0.051| 1.477| 1.677|
| No Cultivated Land                            | 4,130                | 0.053  | 0.009| 0.034| 0.072|
| 0–2 ha of cultivated land                     | 4,130                | 0.565  | 0.019| 0.528| 0.602|
| 2–5 ha of cultivated land                     | 4,130                | 0.296  | 0.015| 0.267| 0.326|
| >5 ha of cultivated land                      | 4,130                | 0.086  | 0.010| 0.066| 0.105|
| Household paid loans to bank or family friends in past year | 4,130                | 0.023  | 0.005| 0.013| 0.032|
| Public Transportation to Regional HQ in EA    | 157                  | 0.756  | 0.038| 0.682| 0.830|
| Electricity in the EA                          | 157                  | 0.238  | 0.037| 0.165| 0.311|
| Mobile Signal in the EA                       | 157                  | 0.821  | 0.034| 0.753| 0.889|
| Internet in the EA                            | 157                  | 0.086  | 0.024| 0.039| 0.133|
| Bank in the EA                                 | 157                  | 0.024  | 0.012| 0.000| 0.047|
| Cooperative Primary Society in the EA          | 157                  | 0.381  | 0.042| 0.298| 0.464|
| Informal Financial Service in the EA           | 157                  | 0.450  | 0.043| 0.366| 0.535|
| Major Employer (that is business, factory) in the EA | 157                  | 0.161  | 0.032| 0.098| 0.225|
| Weekly Market in the EA                       | 157                  | 0.295  | 0.040| 0.217| 0.373|

*Notes:* The estimates account for survey sampling design.  
*Source:* Authors’ calculation using data of 2012 HBS.

### Table A2. Summary statistics of main variables of 2012 HBS used in the regression – urban

| Variable                                      | Number of HHs or EAs | Means  | SE   | LL   | UL   |
|-----------------------------------------------|----------------------|--------|------|------|------|
| Youths headed (between ages 15–34)            | 6,056                | 0.343  | 0.010| 0.324| 0.362|
| Female headed                                 | 6,056                | 0.249  | 0.009| 0.231| 0.267|
| No Primary Education                          | 6,056                | 0.084  | 0.007| 0.071| 0.098|
| Completed Primary Education                   | 6,056                | 0.632  | 0.014| 0.605| 0.659|
| Completed Secondary Education or more         | 6,056                | 0.283  | 0.017| 0.251| 0.316|
| Number of Youth (15–34)                       | 6,056                | 1.747  | 0.042| 1.663| 1.831|
| No Cultivated Land                            | 6,056                | 0.702  | 0.034| 0.635| 0.769|
| 0–2 ha of cultivated land                     | 6,056                | 0.217  | 0.024| 0.170| 0.265|
| 2–5 ha of cultivated land                     | 6,056                | 0.054  | 0.010| 0.034| 0.074|
| >5 ha of cultivated land                      | 6,056                | 0.026  | 0.011| 0.004| 0.049|
| Household paid loans to bank or family friends in past year | 6,056                | 0.039  | 0.006| 0.027| 0.052|
| Public Transportation to Regional HQ in EA    | 223                  | 0.943  | 0.025| 0.894| 0.992|
| Electricity in the EA                          | 223                  | 0.852  | 0.044| 0.764| 0.939|
| Mobile Signal in the EA                       | 223                  | 0.846  | 0.031| 0.786| 0.907|
| Internet in the EA                            | 223                  | 0.292  | 0.037| 0.219| 0.366|
| Bank in the EA                                 | 223                  | 0.155  | 0.024| 0.108| 0.203|
| Cooperative Primary Society in the EA          | 223                  | 0.188  | 0.033| 0.124| 0.252|
| Informal Financial Service in the EA           | 223                  | 0.650  | 0.040| 0.572| 0.729|
| Major Employer (i.e. business, factory) in the EA | 223                  | 0.416  | 0.044| 0.329| 0.503|
| Weekly Market in the EA                       | 223                  | 0.207  | 0.039| 0.129| 0.284|

*Notes:* The estimates account for survey sampling design.  
*Source:* Authors’ calculation using data of 2012 HBS.