G|L|M LIC
Growth and Labour Markets in
Low Income Countries Programme

WP

GLM|LIC Working Paper No. 23 | February 2017

Out with the Old and Unproductive, in with the New and Similarly Unproductive

Microenterprise Dynamics in a Growing Low-Income Economy

Brian McCaig (Wilfrid Laurier University)
Nina Pavcnik (Dartmouth College BREAD, CEPR, IZA, and NBER)

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

UKaid
from the British people

GLM|LIC Working Paper No. 23 | February 2017

Out with the Old and Unproductive, in with the New and Similarly Unproductive

Microenterprise Dynamics in a Growing Low-Income Economy

Brian McCaig (Wilfrid Laurier University)
Nina Pavcnik (Dartmouth College BREAD, CEPR, IZA, and NBER)

GLM|LIC
c/o IZA – Institute of Labor Economics
Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany
Phone: +49-228-3894-0
Fax: +49-228-3894-510
Email: glm-review@iza.org

I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

UKaid
from the British people

GLM|LIC Working Paper No. 23 | February 2017

ABSTRACT

Out with the Old and Unproductive, in with the New and Similarly Unproductive

Microenterprise Dynamics in a Growing Low-Income Economy*

Non-farm microenterprises employ a large share of the overall and non-agricultural workforce in low-income counties. Yet, the lack of large-scale, nationally representative data on microenterprises often precludes a systematic study of their business dynamics. We examine the dynamics of non-farm microenterprises, including entry, exit, hiring, and transitions to the formal sector, with nationally representative longitudinal household surveys in Vietnam. We find that older business do not perform substantially better than their younger counterparts. Business exit is prevalent and inversely related to business performance. Nevertheless, poorly performing exiting firms are replaced by equally poorly performing entrants, which have little hope of growing into a successful business. Incumbents that are more successful initially are more likely to start hiring workers, increase their workforce, and the most successful transition to become a formal enterprise. However, incumbents that undertake these activities are very rare. Our analysis also highlights the strong correlation of characteristics of a key production input – the owner – with business entry and exit, revenue growth and the ability of a microenterprise to transition to a hiring business. The owner's education and past employment experience play a particularly important role. However, education levels of microenterprise owners, including those that hire, are low, especially relative to wage earners in the formal sector. Overall, our analysis suggests that for most microenterprises in Vietnam, there is little potential for sustained growth and employment expansion.

JEL Classification:
O17, O12, J46, L11

Keywords:
firm dynamics, firm entry, microenterprises, entrepreneurship, informality, household business, Vietnam, resource allocation

Corresponding author:
Nina Pavcnik
Department of Economics
Dartmouth College
6106 Rockefeller Center
Hanover, NH 03755
USA
E-mail: Nina.Pavcnik@Dartmouth.edu

* We thank Hanh Nyugen, Vu Huong Dat, Katherine Ziomek, and Zhecheng (Iraday) Yao for excellent research assistance. We thank participants at the Growth and Labor Markets in Low Income Countries: Program Network Meeting in Lima, Peru, the 1st IZA/DFID GLM-LIC Research Conference on New Research on Labor Markets in Low-Income Countries conference at the World Bank, and The Centre for Analysis and Forecasting, Hanoi, Vietnam for comments. This document is an output from a project funded by the UK Department for International Development (DFID) and the Institute for the Study of Labor (IZA) for the benefit of developing countries. The views expressed are not necessarily those of DFID or IZA.

1. Introduction

Non-farm microenterprises employ a large share of the overall and non-agricultural workforce in low-income counties (Banerjee and Duflo (2007), Gollin (2008), de Mel, McKenzie, and Woodruff (2010), La Porta and Shleifer (2008, 2014), Hsieh and Olken (2014)).[1] These businesses are often small, informal, and operated by a single individual. They account for a large share of employment in low-income countries, so that their performance, dynamics, and likelihood of transitions to hiring businesses and to the formal sector are key to understanding the generation of jobs and aggregate labor productivity in low-income countries. In addition, the allocation of resources and productivity differences between microenterprises and larger firms play an important role in explaining the aggregate labor productivity gaps between low- and high-income countries (Hsieh and Klenow (2009), Restuccia and Rogerson (2008), Hsieh and Olken (2014)).

Despite their dominance in low-income countries, systematic, nationally representative empirical evidence on the performance and dynamics of microenterprises in low-income countries is scant (Woodruff (2012)). In this paper, we examine the dynamics of non-farm microenterprises, including their entry and exit, hiring and transitions to the formal sector in a growing low-income country. We focus on Vietnam, where non-farm microenterprises account for 60 percent of the non-agricultural workforce. We use household-level data from four rounds of the Vietnam Household Living Standard Survey (VHLSS), which spans almost a decade of growth, to examine the dynamics of non-farm microenterprises.[2]

Challenges in examining the dynamics of microenterprises and transitions of businesses over long periods are well known (de Mel, McKenzie, Woodruff (2010, 2013), Woodruff (2012), Pages (2012)). As a result, much of what we know about the dynamics of microenterprises and their owners stems from either labor force surveys from urban areas in middle-income countries in Latin America (see a survey by Maloney (2004), Mondragón-Vélez and Peña (2010)), Townsend Thai

[1] For example, informal household-run businesses account for 80 percent of manufacturing employment in India in the 1990s (Nataraj (2011), Hsieh and Klenow (2009, 2011)) and 58 percent of manufacturing employment in Vietnam in 1999 (McCaig and Pavcnik (2015)).

[2] As explained in Section 3, the business module in the VHLSS covers all non-agricultural business activities conducted by households. Thus, it covers private businesses operating in the informal household business sector as well as those registered as a private enterprise and operating in the formal enterprise sector. However, due to the vast number of household businesses relative to private enterprises, when referring to businesses observed in the VHLSSs we often refer to them solely as household businesses even though the datasets include some private enterprises.

Project[3], and recent randomized control trials (henceforth, RCTs). The latter literature has tested the impact of a wide range of interventions, including microcredit provision, business training, relaxing entry conditions, and easing formalization requirements, on microenterprise performance. These studies have made important advances in our understanding of the impact of supply side constraints on microenterprise growth and job creation, in general finding relatively modest effects of these interventions on microenterprise performance.[4] Our paper complements the existing line of work and overcomes some of the data challenges.

First, because most microenterprises are not registered, they are usually not included in the administrative data collected by the government or census or survey data that covers registered enterprises or enterprises above a certain employment cutoff. We rely on a sample of microenterprises from a nationally representative household survey, which includes a module on non-farm businesses owned by a household. Each survey round contains information on about 20,000 to 34,000 non-farm businesses. This enables us to document the dynamics of microenterprises owned by households from a nationally representative sample in a low-income setting rather than smaller and potentially a more selective samples collected for a RCT.[5]

Second, the longitudinal dimension of our data follows a household and its businesses for a period up to four years. Longitudinal data on businesses owned by households that spans 4 years is very rare in low-income country settings (Woodruff (2007, 2012)), so the data provide a unique opportunity to study the dynamics of household-owned businesses.[6] The longitudinal dimension provides a representative sample of the population of households who could potentially choose to start a business, an important piece of information not available in usual firm-level data sets or in RCT samples. This enables us to document systematically the correlates of household business formation. Moreover, the household panel enables us to track businesses over time for a period of up to four years and create a panel of household businesses. This enables us to study systematically

[3] See, for example, Paulson and Townsend (2004), Paulson, Townsend, and Karaivanov (2006), (Kaboski and Townsend (2011, 2012), and Samphantharak and Townsend (2012).

[4] See surveys of this literature by Banerjee (2013) and Banerjee, Karlan and Zinman (2015) for the impacts of microcredit, Bruhn and McKenzie (2014) for registration, and McKenzie and Woodruff (2013) for business training.

[5] See McKenzie and Woodruff (2013, 2014) for a discussion of issues related to small and selective samples used in RCTs.

[6] An exception is data from the Townsend Thai Project. These studies use detailed observational data on a panel of households over a period much longer than is typically done. See Paulson and Townsend (2004), Paulson, Townsend, and Karaivanov (2006), Kaboski and Townsend (2011, 2012), and Samphantharak and Townsend (2012).

the dynamics of microenterprises over a longer period than is usually possible. We focus on the determinants of entry and exit of household businesses and the dynamics of incumbent businesses. Finally, the VHLSS includes data on all businesses owned by a household, regardless of their formal registration status. As a result, we can also trace successful businesses and examine transitions of household businesses to the formal sector.

Third, we document the relationship between these dynamic patterns and characteristics of the business as well as its key input, the owner/operator/manager. Most household businesses have only one worker, the owner, which is a key production input. This person owns the business, works at the business, and manages its daily operations. We examine how characteristics of the business owner/manager shape business performance. The multi-purpose nature of the VHLSS enables us to link the businesses surveyed in the business module to information about the owner from the modules of the survey related to household demographics, education, and employment. This enables us to examine selection into the non-farm microenterprises in a systematic fashion (see de Mel, McKenzie, and Woodruff (2010)). In particular, using the nationally representative labor force module, we can compare education (and other characteristics) of non-farm business owners to characteristics of individuals in other economic activities. Moreover, we use the detailed information on owners of the business to examine the relationship between the characteristics of the owner and business performance. In particular, we can examine how owner characteristics affect the ability of a business to hire workers, to transition from a non-hiring to a hiring business, and to start directly as a hiring business. This is important for understanding the dynamics of job formation in a low-income country. Finally, given the configuration of the panel datasets, we can additionally evaluate the role of an owner's job prior to commencing the microenterprise. This provides us with important information on whether the most successful micro entrepreneurs possessed valuable prior expertise, such as, for example, exposure to new technology and valuable market knowledge in the foreign-invested sector.[7]

Our analysis yields several interesting insights into the dynamics of household businesses. We find that older business do not perform substantially better than their younger counterparts, which is consistent with patterns observed among formal firms in countries such as India and Mexico (Hsieh and Klenow (2011)). The lack of firm growth with age reflects two underlying forces.

[7]Case studies from Bangladesh apparel industry suggest such exposure plays an important role.

First, the flat age profile of businesses is not due to lack of exit of less efficient businesses: business exit is prevalent and inversely related to business performance. However, the entering firms do not appear to perform better than the exiting firms do. Consequently, poorly performing exiting firms are replaced by equally poorly performing entrants, which have little hope of growing into a successful business. Second, initially more successful household businesses display or eventually adapt more of the characteristics associated with firms in the formal enterprise sector. They are more likely to start hiring workers, increase their workforce and some of them even register as a formal enterprise. However, businesses that undertake these activities are rare. For example, only about 10% of firms hire workers, and of the incumbents that do not hire workers initially, less than 5% hire workers in the subsequent period. Transitions of household businesses to formal private enterprises are even rarer. Less than 3% of businesses are registered as an enterprise and among the incumbents initially not registered as an enterprise, only 1.7% of them make this transition in the subsequent period. Our evidence, based on a nationally representative sample, is thus consistent with evidence from several randomized control trials (de Mel, McKenzie, and Woodruff (2010), Schoar (2010)).

Our analysis also highlights important differences in characteristics of the key production input, the owner/manager, across businesses as these characteristics strongly correlate with business performance. Business growth is positively related to the owner's education, but individuals with low levels of education run many of these businesses. The most educated individuals are not business owners, but wageworkers in the formal enterprise sector. Household businesses started by owners who previously worked as an employee in the enterprise sector are relatively rare. Most owners entering or exiting the household business sector are doing so from or to self-employment in agriculture and have low levels of education, especially relative to wageworkers in the formal enterprise sector. Business owners have levels of education comparable to wageworkers in the informal sector. Second, consistent with de Mel, McKenzie, and Woodruff (2010), we find that owners that hire workers (i.e., employers) tend to be much better educated than owners that do not hire (own-account owners). Large education gaps suggests that it might be difficult for a non-hiring business to transition to a hiring one, potentially explaining the lack of transitions to hiring businesses discussed above. In fact, the longitudinal dimension of the nationally representative data highlights that transitions from own-account worker to employer are very rare: less than 6% of incumbent businesses that do not initially hire workers, do so in the subsequent

period and transitions are more likely to occur in businesses run by better-educated owners. Likewise, better-educated owners are more likely to start a business that starts as a hiring business. Finally, even owners that run a business that hires workers (employers) have substantially lower levels of education than wageworkers in the formal enterprise sector. Overall, this analysis emphasizes the strong correlation of an owner's education with business performance and highlights the relatively low levels of education among non-farm business owners, especially relative to the wage earners in the formal sector.

This analysis suggests that for most businesses owned by households, there is little hope of growth and employment expansion. Expanding business that start informal and transition to the formal sector and hire workers are rare, even in a low-income country setting during a period of fast aggregate output growth. This has important consequences for understanding the dynamics of aggregate share of informal employment in low-income countries. During the past decade, Vietnam experienced a large reduction in the share of employment in informal firms (McCaig and Pavcnik (2015)). Our analysis suggests that formalization of previously informal microenterprises does not drive this decline. Instead, it is happening either through the expansion of existing formal firms in the enterprise sector and/or direct entry of firms to the enterprise sector. Understanding the drivers of expansion of the formal sector might thus be a fruitful future area of research (see also Hsieh and Olken (2014)). Recent work emphasizes the importance of international trade and deregulation of small-scale manufacturing in this process (McCaig and Pavcnik (2014), Martin, Nataraj, and Harrison (2014)).

Our study complements the literature on microenterprises using randomized control trials. These studies have made important advances in our understanding of the impact of supply side interventions such as microcredit provision, business training, relaxing entry conditions, and easing formalization requirements on microenterprise growth and job creation.[8] They in general find relatively modest effects of these interventions on microenterprise performance. This is consistent with our findings that indicate that very few microenterprise owners possess the characteristics needed to run a successful business and that business expansion, hiring of workers, and transition to the formal sector are very rare.

[8] See summaries by Banerjee (2013) and Banerjee, Karlan and Zinman (2015) for the impacts of microcredit, Bruhn and McKenzie (2014) for registration, and McKenzie and Woodruff (2013) for business training.

At the same time, recent surveys of this literature have highlighted some of the data challenges inherent in the existing RCT work on microenterprises. These challenges include short time periods, relatively small samples of very heterogeneous firms, and a lack of information about the markets in which these firms operate (see Banerjee, Karlan and Zinman (2015), Bruhn and McKenzie (2014), McKenzie and Woodruff (2013).[9] While our analysis is descriptive in nature, the large-scale sample, which is based on nationally representative data of households across all geographic regions (urban and rural) and industries, enables us to systematically documents the underlying dynamics of entering, exiting, and surviving microenterprises and compares the owners of microenterprises to individuals in other economic activities. Importantly, we can follow businesses, their owners, and their associated households over a period of up to four years. Further, the ability to follow associated households allows us to investigate characteristics and activities of individuals before and after they operate a business. Since several low- and middle-income countries have household surveys similar to those in Vietnam, which are based on the World Bank's Living Standards Measurement Study program, we see this as a potentially rich source of data for future analysis, complementing the samples used in RCT studies. Indeed, perhaps future RCTs can take advantage of this existing survey infrastructure.

The paper proceeds as follows. Section 2 defines a household business and summarizes related literature. Section 3 describes the data. Section 4 analyses facts about the dynamics of businesses owned by the household. Section 5 examines the role of business owners/managers, an important production input, in business performance. Section 6 concludes.

2. Background and related literature on microenterprises in Vietnam

We first discuss the empirical definitions of microenterprises and informality used in the existing literature. Next, we overview the definition of household businesses in Vietnam and discuss how the definition of informality based on household businesses relates to other definitions used in the literature.

2.1. Definition of a Household Business

[9] See McKenzie and Woodruff (2014), de Mel, McKenzie, and Woodruff (2010, 2013), Bruhn and McKenzie (2014), and Banerjee, Karlan, and Zinman (2015) for discussion.

Informality is a multi-dimensional concept (see Goldberg and Pavcnik (2004, 2007), La Porta and Shleifer (2008, 2014) for surveys). In this paper, we focus on one dimension of informality, namely a distinction on whether a firm is a household business or an enterprise. In Vietnam all state, foreign and collective businesses are legally required to register as enterprises under Vietnam's Enterprise Law.[10] However, private businesses can legally operate in three different modes: a household business without a license, a household business with a license, and a private enterprise.[11] Rules exist describing in which of the three modes the business should be operating, but many household businesses seem unaware of the legal requirements to register their household business despite almost all of them meeting the legal requirements to do so based on their income (Cling, Razafindrakoto, and Roubaud (2012)).[12,13] Thus, any private business which is not registered as an enterprise is, broadly speaking, considered to be a household business. Being a registered private enterprise is commonly used as a measure of a formal business (La Porta and Shleifer (2008)). McCaig and Pavcnik (2014) provide detailed discussion of the distinction between a household business and private enterprise. Among others, they report, "registered enterprises are required by the Enterprise Law to follow formal accounting standards and to report comprehensive information about their financial position, including information on their workforce. Consequently, as in other low-income countries, in Vietnam workers in the enterprise sector are captured in the conventional firm-level datasets based on administrative records covering the formal sector,

[10] During our study period, the relevant version of the Enterprise Law is the Law on Enterprises passed in 1999.
[11] See Decree No. 02/2000/ND-CP and Decree No. 109/2004/ND-CP.
[12] Household businesses operating without a license are not necessarily doing so to avoid detection by tax or labor authorities as businesses engaged in salt-making, as street vendors, or as service providers with low incomes are not required to register. See Decree No. 109/2004/ND-CP.
[13] Various studies discuss the benefits and costs of a firm being an enterprise as opposed to a household business. For example, Malesky and Taussig (2009) report that enterprises, relative to household businesses, have easier access to export licenses, customs certificates, opportunities to bid on government contracts, the right to open branches and to operate outside their home district. At the same time, running an enterprise (as opposed to a household business) entails the registration cost and more rigorous accounting. Taussig and Hang (2004) reports benefits of being an enterprise (relative to household business) as greater ability to trade beyond home district, ability to expand, value added tax receipts, legal ability to establish branch locations, a stamp for making transactions more official, more predictable, law based interactions with government, ability to access equity for limited and joint stock companies, and greater access to government investment incentives. Costs of formalization include registration costs, annual registration fee, certified chief accountant, greater reporting requirements, potential for increased attention from local authorities, and potential for increased taxes with movement from lump sum to standard tax calculations. They also report that many laws governing household businesses are the same as those for sole proprietorships, the simplest form of a company (CIEM, *Assessment of the Strengths and Weaknesses of the Enterprise Law: Recommendations for Amendments and Additions (Draft),* 2004, p. 62). The information on the costs of registering as a private enterprise in Vietnam is summarized by the World Bank's Doing Business Survey.

whereas workers in the household business sector are not (page 10)." Our project focuses on household businesses and private enterprises owned by households and relies on data that includes both of these establishments.[14]

In addition, in Vietnam another measure of formalization is whether a household business has a business license. As a result, our analysis also uses holding a business license as a marker of the extent to which a business is integrated into the "formal" economy. Household businesses operating with a business license would not be considered as formal enterprises by the above definition because they are not registered as private enterprises. However, obtaining a business license is considered as a first move toward formality (Cling, Razafindrakoto, and Roubaud (2010, 2012)).[15] Information on the costs of obtaining a business license is not readily available, but costs do not seem to be a significant barrier. Only a small percentage of household businesses without a license consider the lack of license being due to the expense or complicated nature of the licensing process (World Bank (2010)) and World Bank (2009) reports that few households experienced difficulty registering a household business.[16] Among households businesses without a license, the vast majority reports they are not registered either because it is not compulsory or because they don't know if they need to register (World Bank (2010)). Household businesses that hold a license report the main advantage of having a license as less corruption, followed by better access to market places, and easier loan access (World Bank (2010)). One possible explanation for not obtaining a business license is that a requirement of being licensed is to pay taxes. However, Cling et al. (2012) find that some unlicensed businesses report paying taxes and most household businesses also make additional payments to public officials. Therefore avoiding taxes by being

[14] Most of the studies using the 1993 and 1998 Vietnam Living Standards Surveys and the 2002 and 2004 Vietnam Household Living Standards Surveys simply refer to the businesses included in the datasets as non-farm household enterprises (NFHE). In 1993, Vietnam did not distinguish between household businesses and private enterprises. This began sometime later, but the distinction did exist by the 1998 survey. Neither the 2002 nor the 2004 survey distinguishes between household businesses and private enterprises. This is probably why researchers use the term non-farm household enterprise since it incorporates both types of businesses covered in the business modules. However, data from the 2006 and 2008 VHLSSs, which distinguishes between household businesses and private enterprises, suggest that even in 2006 and 2008 a very small share of private businesses, less than 3%, are private enterprises.

[15] In Vietnam, these businesses are often referred to as formal and informal household businesses respectively.

[16] World Bank (2010) is based on a survey of licensed and unlicensed household businesses operating in Hanoi and Ho Chi Minh City, while World Bank (2009) is based on data from the 2008 VHLSS. A web site aimed at the business community (http://www.vietnam-briefing.com/news/vietnams-taxes-business.html/) suggests that the fee varies by monthly income of the business from 50,000 to 1,000,000 dong per year, but the site does not list the original source for this information nor the relevant period.

unlicensed may not lead to an overall decrease in payments to government agencies. Despite imperfect adherence to the law, obtaining a business license is considered as a first move toward formality (Cling et al. (2010, 2012)).

Several earlier studies have thus focused on this definition of household businesses in Vietnam (see, for example, Vijverberg et al. (2006)). Cling et al. (2010, 2012) used an alternative definition in two studies on the informal sector in Vietnam.[17] Following the official ILO definitions, these two studies define the informal sector "as all private unincorporated enterprises that produce at least some of their goods and services for sale or barter, are not registered (no business license) and are engaged in non-agricultural activities," (Cling et al. (2010), p. 6). Hence, this definition corresponds to unlicensed household businesses. This study further defines informal employment "as employment with no social security (social insurance)" (Cling et al. (2010), p. 6).

Our definition of informality based on household businesses versus firms in the enterprise sector falls somewhere between the two definitions above. It is broader than the concept of the informal sector used in existing literature on Vietnam, but likely to be more restrictive than the concept of informal employment, which also includes workers in the enterprise sector who do not receive social insurance.[18] Given the nature of our data, we can perfectly replicate the informal sector definition used in Cling et al. (2010, 2012), but we cannot do so for the informal employment definition.[19]

3. Data and business characteristics

3.1 Data Description

We use data from the 2002, 2004, 2006, and 2008 Vietnam Household Living Standard Surveys (VHLSS), which the General Statistics Office (GSO) of Vietnam conducted. An important advantage of these household surveys is that they are nationally representative and contain a detailed business module on non-farm businesses owned and run by the households. These businesses can be either household businesses or private enterprises.[20] Our data thus enables us to

[17] Cling et al. (2010) was commissioned by the ILO.
[18] In Vietnam, enterprises are required to make social security contributions on behalf of their workers whereas household businesses may voluntarily make such contributions.
[19] Beginning in 2007 the GSO began conducting a nationally representative labor force survey. The questions were designed to be consistent with the concepts of the informal sector and informal employment as defined by the ILO and these surveys were the basis of analysis for Cling et al. (2010).
[20] Analysis of the 2006 and 2008 VHLSSs suggest that less than 3% of the businesses are private enterprises.

study a representative sample of businesses owned by households, which is rare since most of these businesses are usually not covered in the more widely available firm-level data. The later tend to focus only on registered (or formal) firms or formal firms above a certain employment size cut off. They thus preclude the study of dynamics of small and/or unregistered firms, which account for the largest share of employment and businesses in low-income countries.

Each survey round contains information on about 20,000 to 34,000 businesses. The business module collects information on whether a household operates a business, the industry in which the business operates, the number of months it operated during the past 12 months, revenue, the wage bill, other expenditures (materials, repairs, water, etc.), whether the business has a license, and who is the most knowledgeable person (hereafter referred to as the owner/manager).[21] The survey also contains information on the number of workers (available in 2004, 2006, and 2008), the year the business started (available in 2004, 2006, and 2008), and whether the business is registered as an enterprise as opposed to a household business (available in 2006 and 2008).[22]

Importantly, the survey contains two- and four-year panel subcomponents at the household level. Longitudinal dimension of data on businesses owned by households that spans four years is very rare in low-income country settings (Woodruff (2007, 2012)), so the data provide a unique opportunity to study the dynamics of household-owned businesses. We use the longitudinal dimension of the data to create a panel of businesses that spans a period of two to four years. In particular, between each successive survey there is a panel of approximately 20,000 households. Additionally, a subsample of the panel households was surveyed for a third round. Thus, we have three two-year panels of approximately 20,000 households (2002/04, 2004/06, and 2006/08) and two four-year panels of approximately 10,000 households (2002/04/06 and 2004/06/08). Longitudinal data is useful for two reasons. First, it provides for a representative sample of households, who could potentially choose to start a business, an important piece of information not available in usual firm-level data sets. This enables us to study the determinants of household

[21] The information on the most knowledgeable person is not available for 2004 VHLSS. However, we have designed an algorithm to predict the most knowledgeable person. See the data appendix for further details in the procedure.

[22] The business modules of the 2002 and 2004 VHLSS do not distinguish whether a business run by the household is a household business or a private enterprise. In 2004 we can use information on whether the business manager/owner reports working in the private sector (as opposed to the household business sector) to gauge the prevalence of businesses that are private enterprises. Only 1.5 percent of panel businesses could be considered private enterprises by this definition. The share would likely be even smaller (or similar) for 2002.

business formation. Second, household panel data enables us to track household businesses over time and create a panel of household businesses.[23] The four-year panel contains about 6,500 businesses. We create a panel of household businesses by taking advantage of the information on the number of household businesses a household operates and the information about the business that is not likely to change in a short period.

In particular, we use information on the industry and manager of the business to match businesses between the surveys.[24] We construct a business panel by first matching businesses over time within a household by industry and manager. Subsequently, among remaining businesses we match by either manager or industry within the household. We provide full details of the procedure in the Data Appendix. We match 14,229 businesses by industry and manager, 1,660 businesses by industry only, and 4,045 businesses by manager only leading to a panel of 19,934 businesses. This represents 92 percent of all possible businesses that could be matched over time within panel households.

Finally, we can link the household business module to the information about its key input, the owner/manager. We obtain information on the business owner by linking information on household businesses to information on the owner from the modules of the survey related to household demographics, education, health, employment. For each working individual, the employment module collects information on the industry of affiliation, occupation, ownership category (self-employed, working for other households, state-owned economic sector, collective economic sector, private economic sector, foreign-invested sector), hours worked during the past 12 months, number of years doing the job, and wages and other benefits. Demographic modules contain information on the owner's age, gender, ethnic minority status, completed level of education, and location. Note that the VHLSS also includes a longitudinal dimension at the individual-level among the corresponding members of households included in a household-level

[23] The VHLSSs feature a rotating panel by enumeration area. Thus, not all enumerations areas surveyed in 2002 were intended to be resurveyed in 2004. This accounts for why the number of panel households is noticeably lower than the total number of households surveyed in the 2004 VHLSS, as only about half of the enumeration areas surveyed in 2004 were surveyed in 2002.

[24] Because the 2004 VHLSS did not report the manager of the business and thus before we can match businesses over time we begin by predicting the manager of the business by matching information reported about the business with information reported in the labor module about workers within the household. A full description of the process is available in the Data Appendix. We test the procedure using the 2006 VHLSS, which contained the same business information as the 2004 VHLSS but also reported the manager and find that our procedure correctly predicts the manager for 92.4 percent of businesses in the 2006 VHLSS. Thus, we feel very confident about our ability to predict accurately the manager for businesses reported in the 2004 VHLSS.

panel data.[25] The longitudinal dimension at the individual level enables us to examine economic activities of business owners prior to starting a business and after closing down a business. We focus on individuals ages 20 to 64. Overall, we use the detailed information on owners of the business to examine the relationship between the characteristics of the owner and business performance.

3.2 Business Characteristics

Table 1a provides information about the number of households and number of businesses in each survey. The sample size ranges from 46 to 74 thousand household.[26] 37 percent of households own a non-farm business, with almost 80% of these households owning only one business, and 19% of these households owning two businesses. Households with more than two businesses are rare.

Table 1b provides summary statistics about businesses and their owners. The surveys include information about 20,000 to 34,000 businesses.[27] About 70% of non-farm businesses operate in services (i.e. tertiary activities), while manufacturing accounts for the vast majority of the remaining household businesses. Within manufacturing, household businesses are most common in food and beverage production, wood processing, clothing and apparel, furniture manufacturing, and the textile industry. Most businesses are in rural areas, although the share of businesses located in urban areas has increased from 33 to 40%. 19 to 27% of them have a business license, a first step toward formality, but only 2.3% are registered as a private enterprise (a formal business). The average business operates 10 months a year.

Household businesses employ on average only 1.6 employees, in contrast to 63.8 workers per firm in the enterprise sector (McCaig and Pavcnik (2014)). 65% of businesses report only one worker, 25% have two workers, with less than 1% employing ten or more workers. Some businesses

[25] The surveys do not follow individuals that left a household, although the datasets often provide some basic information of why they left (i.e., for marriage, for work, etc.) and where they move.

[26] The sample size is largest in 2002. The decline in the sample size between the 2002 and later rounds of the VHLSSs is primarily due to a reduction in the number of households surveyed within an enumeration area. For example, the 2002 VHLSS surveyed households within 3,001 enumeration areas averaging approximately 25 households per enumeration area. The 2004 VHLSS surveyed 3,062 enumeration areas, but only 15 households per enumeration area.

[27] Table 1a reports the number of households operating a particular number of businesses. Thus, the total number of businesses in Table 1b, can be obtained from Table 1a by summing over the product of the number of households times the number of businesses being operated by those households.

report a large number of employees (the largest has 300 workers), but such large businesses are extremely rare. About 10% of businesses report hiring a worker. The low number of employees per household business and lack of hiring is consistent with employment patterns in informal businesses from other less developed countries surveyed in Woodruff (2007) and Schoar (2010). Household businesses have lower labor productivity than enterprises (McCaig and Pavcnik (2014)), which confirms findings of other studies (La Porta and Shleifer (2008)).

The bottom panel of Table 1b summarizes characteristics of household business owners/managers, a key input into production and running of the business. About 75% of businesses are operated by an owner as their primary job. Females operate almost 60% of businesses. The average business has an owner that is a 40-year old, with less than 8 years of complete education. Ethnic minorities run only 5% of businesses.

Household businesses exhibit a large degree of heterogeneity in performance. Figure 1 plots the density of log revenues for household businesses and shows large differences among household businesses. This is consistent with the literature on firms operating in the formal sector, which documents a high degree of heterogeneity in underlying performance (see Melitz and Redding (2014) for a survey). However, these studies do not capture household businesses, which suggest that they underestimate the degree of productivity heterogeneity. In fact, Vietnamese household businesses tend to be smaller (as measured by revenue or employment) and have lower labor productivity than firms in the enterprise sector (McCaig and Pavcnik (2014)). Although many studies ignore this left tail of the firm size and productivity distribution, it is potentially crucial to understanding the dynamics of aggregate productivity in low-income countries. A key feature of the firm size distribution in these countries are firms with one employee, the owner, such as businesses owned by households in Vietnam (Hsieh and Olken (2014), La Porta and Shleifer (2008, 2014)). In Vietnam, non-farm household businesses employed 60% of the workforce outside of agriculture in 1999 and 58% of the workforce in manufacturing (McCaig and Pavcnik (2015)). As a result, understanding the dynamics of their performance is fundamental for understanding aggregate labor productivity in Vietnam. We turn to that next.

4. The dynamics of household business performance

4.1. Age profile of household businesses

Market selection forces would suggest that only the better household businesses should survive over time, yielding a positive relationship between household business age and revenue. Hsieh and Klenow (2011) report such a profile among formal firms in the U.S. Figure 2 plots the distribution of household businesses by age in 2008. Most businesses in Vietnam are very young (43% are 5 years or younger) and 24% are older than 10 years. Interestingly, older businesses do not tend to perform better than younger ones. Figure 3a plots the average number of workers employed by businesses by age in 2004, 2006, and 2008.[28] If older businesses outperformed younger businesses, one would expect substantially larger average employment in older firms. However, the relationship between employment and age is fairly flat, so that older businesses do not have higher employment than younger ones. Figure 3b performs similar analysis for revenue in 2008, with 95% confidence bands and suggests a flat age profile for revenue: older firms do not have statistically different revenue than younger firms. Finally, Figure 3c plots average revenue per worker by age of a business in 2008, with 95% confidence bands. If better firms were more likely to survive over time, we would expect older firms to have higher labor productivity than younger ones. Yet, older firms are not statistically significantly more productive than younger ones. One potential concern about the analysis is that better performing firms would formalize, thus no longer operating as household businesses. However, as discussed in section 3, our data includes businesses owned by households, regardless of their registration status, so that the more successful firms continue to be in the sample if they transition to private enterprise.[29]

The age profile patterns noted in Vietnam in Figures 3a-c are consistent with patterns observed among formal firms in countries such as India and Mexico (Hsieh and Klenow (2011)). Nonetheless, they differ from the age profile of firms in the U.S., where older firms tend to be substantially bigger in terms of employment, market share, and labor productivity than younger firms are. Our analysis suggests that the forces of market selection do not appear to eliminate less productive firms over time. Alternatively, they suggest that successful firms cannot grow over time, perhaps due to constraints they face.

[28] The average business has operated for 7.7 years. Only 10% of household businesses are more than 20 years old, so we confine the analysis to household businesses that have operated for at most 20 years.

[29] Unreported analysis suggests that there is almost no relationship between business age and the likelihood of enterprise registration. If anything, the probability of being a registered enterprise is decreasing with age, but the difference between a new and 10 year old business is not statistically significant.

4.2. Longitudinal analysis of the dynamics of businesses

The above analysis compares the performance of younger and older businesses at a point in time in a (repeated) cross-section. The longitudinal dimension of the data enables us to further explore why older business do not on average perform better than their younger counterparts do.

One potential explanation for why older firms do not perform better than younger firms could be the lack of exit of less successful firms. However, exit rates are high, with 30 to 35 percent of businesses shut down between two survey rounds. These rates of entry and exit exceed the rates reported for formal firms (Roberts and Tybout (1996)). In addition, Figure 4 plots the relationship between business revenue and the propensity that a firm exits by next survey round. If market selection does not eliminate poor performing businesses, we would find no relationship between business revenue and future exit. However, exit is negatively related to business performance, with the probability of exit being the largest among the poorly performing business. Our evidence is consistent with World Bank (2010), which finds that initially better performing household businesses (as measured by initial revenue, value added, profit, and employment) and older businesses are less likely to exit, and with evidence on the importance of selection for the formal sector firms (Pavcnik (2002), Melitz (2003)).

The question then becomes why does household business performance not improve with age despite selection forces at work? One possibility is that businesses that perform equally poorly (and/or no better than survivors do) replace those that fail. In Figure 5a, we plot real monthly revenue for entering businesses, exiting businesses, and incumbent businesses based on the 2004-2008 panel. We define an entrant as a business that did not operate in 2004, but operated in 2006 and an exiter as a business that did not operate in 2008, but operated in 2006. An incumbent operated in all three years, from 2004-08. The reported revenue series for each group of firms is based on monthly revenue in 2006. Thus, the differences in revenue across the three groups of firms do not arise due to differences in the length of operation in within a year.[30]

Several facts emerge. First, the revenue distribution of entering and exiting firms is shifted to left of the distribution for the incumbent firms, consistent with survival of better performing firms noted in Figure 4. What is perhaps more surprising is the perfect overlap between the revenue distribution of entering and exiting businesses. While poorly performing businesses are

[30] Since most businesses only employ one worker, the owner, there is not much difference between monthly revenue and monthly revenue per worker.

more likely to exit, the entering businesses tend to be equally poorly performing as their exiting counterparts. If we exclude firms that entered and exited during the same period, so that the distribution of entrants is conditional on surviving the first period, entering businesses are slightly better than exiters (see Figure 5b).[31] Nonetheless, the two distributions continue to overlap closely.[32] This evidence suggests that while the entry and exit rates are high, the entering businesses do not perform better than the exiting businesses. One potential explanation for the overlap in the distributions is that households form non-farm businesses to smooth economic shocks (see, for example, Adhvaryu, Kala, Nyshadham (2013)). Establishment of a new business then provides an important coping mechanism for households that lack alternative ways to deal with negative shocks.[33] Others have also proposed that businesses exist because their owners do not have alternative means of employment (La Porta and Shleifer (2008, 2014)). This raises a question about the likelihood of a new entrant growing into a successful business. The comparison of revenue distribution of entering and exiting businesses, which is based on longitudinal data, rather than cross-sectional comparisons based on age of business in earlier work, is consistent with the view that entering businesses tend to be poorly performing. Our analysis suggests that many of them are likely to exit in the near future and the likelihood of them developing into a high-performance business is low. The latter is consistent with a substantial overlap in the distribution of incumbent firms with the distribution of entering and exiting businesses in Figure 5a. While the distribution of the incumbents is clearly shifted to the right of the distribution of exiting and entering firms, the overlap in the distribution is nonetheless large.[34]

We use the longitudinal nature of our data to examine further the dynamics of incumbent firms, including whether incumbents that are initially more successful expand. Figure 6 first plots the relationship between initial monthly revenue of a business and subsequent revenue growth. Initially better performing businesses have lower revenue growth rates. This negative relationship could be due to measurement error. It could also reflect mean reversion.

[31] 16% of the firms in Figure 5a entered and exited during the same period.
[32] In future work, we will also compare the distribution of entrants in 2006 that survive until 2008 to the distribution of firms that exited in 2006 to further examine differences in the evolution of performance of entrants (relative to exiting firms) over time.
[33] In future work we will examine whether the VHLSS allows us to identify which businesses might have been started as a coping mechanism.
[34] In future work we will formally test whether the distributions are statistically distinguishable from each other.

In addition to revenue growth, the existing literature has emphasized hiring of paid workers (outside of the family), increased workforce, and formalization as indicators of business expansion (de Mel, McKenzie, and Woodruff (2010, 2012)). Most of our analysis also focuses on these outcomes. Our longitudinal analysis proceeds as follows. We pool the data from the all the years and in each year focus on incumbent firms that did not hire workers in the previous period. Among the firms without hired workers in period t, we then examine the probability that the firm hires workers in the subsequent survey as a function of its performance (measured by revenue) in previous period. Figure 7a presents the results. It plots the probability of hiring outside labor versus past monthly revenue for businesses that did not hire labor in the previous period. The relationship is strongly positive. The relationship between the probability of hiring workers and performance is flat for low-performing incumbents, but increases with firm past performance for the best preforming firms. As the low propensities on the vertical axis suggest, starting to hire workers is rare. Only 4.7 percent of firms that initially did not hire workers start to hire workers in the subsequent period. We perform similar analysis for all incumbent firms, but use a change in the number of workers as a dependent variable. Figure 7b presents the results and indicates that initially better performing incumbents are more likely to increase their workforce in the subsequent period. The large drop in the right most tail of the distribution is driven by one business that decreased its workforce by 37 workers.

We also consider the relationship between incumbent performance and measures of business formality, namely whether the business has a business license and whether the business is a registered private enterprise. As discussed in section 2, obtaining a business license is the first step toward formalization of a household business in Vietnam. We focus on businesses that do not hold a business license. We then plot the relationship between business revenue and the probability that the business obtains a business license in the subsequent period. As Figure 8a suggests, more successful incumbents are more likely to obtain a business license. Our findings of a positive correlation between business performance and obtaining a business license are consistent with evidence in World Bank (2010). Their study relies on data that distinguishes between licensed

(i.e., formal) and unlicensed (i.e., informal) household businesses based on an alternative survey of Vietnamese household businesses.[35]

Most importantly, we also examine the relationship between initial performance and probability that a firm formalizes, i.e., transitions from being a household business to a private enterprise. The 2006 and 2008 VHLSS data distinguishes between household businesses and private enterprises, enabling us to examine this process. Figure 8b plots the relationship between becoming a private enterprise and past revenue for businesses that were initially not registered (i.e. household businesses). Better performing businesses are more likely to become registered as an enterprise. However, this event is very rare: only 1.7% of businesses transition to become private enterprises and these are heavily concentrated among the best performing businesses. Our evidence on formalization is consistent with the findings on the role of selection into formalization by de Mel, McKenzie and Woodruff (2013), Maloney (2004), Rodgers and Swinnerton (2004), and La Porta and Shleifer (2008). That is, better businesses are more likely to formalize than formalization having a positive impact on performance of businesses.

In sum, our results suggest that the flat age profile of businesses can be explained by the following dynamics of household businesses. Although there is a lot of entry and exit is inversely related to business performance, the entering firms do not appear to perform better than the exiting firms. Consequently, poorly performing exiting firms are replaced by equally poorly performing entrants, who have little hope of growing into a successful business. Related, our analysis suggests that initially incumbents that are more successful are more likely to start hiring workers, increase their workforce, and register as an enterprise. This suggests that initially more successful household businesses display or eventually adapt more of the characteristics associated with more formal interactions with government officials and labor markets and associated with firms in the formal enterprise sector. However, these activities are extremely rare. We find that only 10% of the businesses hire labor, and less than 5% of the incumbents start hiring labor. Transitioning to a private enterprise is even rarer. Less than 3% of businesses are registered as an enterprise, and 1.7% of them make this transition during our sample. This analysis suggests that for

[35] The Household Business and Informal Sector Survey (HB&IS) used in the study is only available for 2007 and 2009 and only covers Hanoi and Ho Chi Minh City. They find that formal household businesses tend to be bigger (in terms of revenue and employment size) and less likely to exit than informal household businesses. Furthermore, informal household businesses that become formal tend to be bigger and more productive than informal household businesses that remain informal.

most businesses, there is little hope of growth and expansion. Businesses that successfully grow, start hiring workers, and transition to the formal sector are rare. These underlying characteristics of household business dynamics, which become apparent through the longitudinal dimension of the data, help explain why the age profile of businesses owned by households is flat.

5. The role of a key input into production: The owner/manager

Most household businesses have one worker, the owner. This person owns the business, works at the business, and manages its daily operations. In this section, we examine how characteristics of the owner/manager shape business performance.

We obtain information on the business owner by linking information on household businesses to information on the owner from the employment and demographic modules of the VHLSS. The sample is restricted to individuals ages 20 to 64. Who is a typical owner of a household business? Recall the bottom panel of Table 1b, which summarizes characteristics of household business owners/managers. About 75% of businesses are operated by an owner as their primary job. Females operate almost 60% of businesses. The average business has an owner that is a 40-year old, with less than 8 years of complete education. Ethnic minorities run only 5% of businesses despite constituting XX% of the workforce.

5.1 Owner Education and Experience

With this in mind, we first focus on the role of owner education. Growing a business is positively correlated with owner's education. Figures 9a and 9b plot the relationship between owner's education and the change in the number of workers (Figure 9a) and revenue growth (Figure 9b).[36] Owner education is strongly positively correlated with both measures of business performance. Businesses run by more highly educated owners experience a greater increase in the number of workers and have higher revenue growth. However, even the most educated owners in our dataset saw their number of workers increase by only 0.15 workers over two years.

Individuals with relatively low levels of education run most household businesses, especially when compared to wage workers in the formal enterprise sector. Using the nationally

[36] Figure 9a is based on the 2004-06 and 2006-08 panels, while Figure 9b is based on all three two-year panels. We cannot use the 2002-04 panel to examine changes in the number of hired workers because the 2002 VHLSS did not ask businesses about the number of workers.

representative labor-force module, we compare average education levels of non-farm business owners to individuals in other economic activities (see Table 2). Household business owners are more educated than self-employed and wage workers in agriculture. However, owners of businesses in the non-farm business sector have substantially lower levels of education than wageworkers in the enterprise sector. Wageworkers in the enterprise sector have, on average, 10.5 years of complete education. In comparison, managers of non-farm businesses that do not hire workers have 7.7 years of education and managers of non-farm businesses that do hire workers have 9 years of education.

We find that owners of business that hire are more educated than owners of business that do not hire. This is consistent with existing studies that suggest microenterprise owners tend to have lower levels of education than businesses that hire workers and comparable education to wage workers (de Mel, McKenzie, Woodruff (2010), Woodruff (2012)). Owners that do not hire have similar education levels as wageworkers in the household business sector. However, even the owners businesses that hire workers tend to be less educated than wageworkers in the enterprise sector.

Overall, this suggests that the average human capital of the owners of household businesses is low and that there are large education gaps between being an owner of a business that does not hire and a business that does hire, which makes movements across these categories less likely. The most educated individuals do not run a non-farm business. Instead, they work as wageworkers in the enterprise sector.[37]

The longitudinal dimension of the VHLSS at the individual level enables us to examine experience in economic activities of business owners prior to starting a business and after closing down a business.[38] Table 3 examines the activities of owners before they start a business and after they close down the businesses. Importantly, the majority of owners that start a business are transitioning from self-employment in agriculture. Note that education levels of owners in agriculture are the lowest among all groups in Table 2. Among the employed, about 10 percent of owners start a business after being employed as a wageworker in the enterprise sector (and 12%

[37] In future work, we will use occupation categories of wageworkers in the enterprise sector to summarize education levels of managers and leaders in SOEs and FDI firms.

[38] The VHLSS includes a longitudinal dimension at the individual level among all members of households that are included in the household panel. The surveys do not follow individuals that left a household, although the datasets often provide some basic information of why they left (i.e., for marriage, for work, etc.) and where they moved.

leave for that sector). Transitions from wage work in the household business sector or self-employment in household sector without the managerial obligations are more common. It is thus not surprising that the mean completed grade of owners of an entering business is 7.8 years. This level of education is similar to the education of owners of non-hiring businesses and wageworkers in the household business sector. Likewise, most owners that shut down a non-farm business transition to self-employment in agriculture. Among the employed, about 12% transition to wage work in the formal enterprise sector. Transitions to wage work in the household business sector or self-employment in household sector without the managerial obligations are more common. The mean grade of owners of exiting businesses is 7.9 years: this level of education is similar to the education of owners of non-hiring businesses, wageworkers in household business sector, and owners of entering businesses.

This analysis suggests that transitions are more likely across employment groups with more comparable educational attainment. Furthermore, low and very comparable levels of education among owners of newly established businesses and owners of businesses that are shutting down might help explain the overlap in performance of entering and exiting businesses in Figures 5a and 5b.

5.2 Who starts hiring workers?

We use the cross-sectional and longitudinal dimensions of the VHLSSs to compare owners of businesses that do not hire (i.e., own-account workers) to owners of businesses that do hire (i.e., employers). An important advantage of our setting is that information on different types of business owners is drawn from a nationally representative household (and individual-level) survey, which captures households and individuals in wage work, owners of businesses that do not hire (i.e., own-account workers), and owners of businesses that hire (i.e., employers). In addition, the longitudinal dimension of the data enables us to examine the transitions between different types of businesses. In particular, we can examine whether and how owner characteristics affect the probability that a business transitions to hiring workers.

We begin by examining the relationship between owner characteristics and the probability of running a business that hires workers (rather than a business that does not hire workers). The sample is restricted to individuals age 20 to 64 inclusive and is based on the primary job of the individual. Business managers are matched to the highest ranking business that the individual

manages if managing more than one business. In particular, we estimate a linear probability model of the indicator that a manager hires workers as a function of the manager's education, gender, age, urban versus rural residence, and ethnic minority status. We also include survey year and province fixed effects, so that we are comparing owners of businesses within the same local labor market at a point in time. Table 4 presents the results for the overall sample in column 1, as well as for rural and urban areas in columns 2 and 3, respectively. Because the results are similar for urban and rural areas, we focus the discussion on the overall pooled sample in column 1.

Note that being an owner of a business that employs worker is rare, even among business owners that are managing their business as a primary activity, as only 12.4% of these owners are employers.[39] In addition, the regression analysis in Table 4 suggests that employers differ from own-account workers in observable dimensions. In particular, relative to own-account workers, employers are more likely to be better educated, male, older (particularly of middle age), not an ethnic minority, and live in urban areas. Finally, while observable characteristics of owners such as education correlate positively with ability to hire, the low R^2 suggests that most of the heterogeneity in being an employer remains unexplained. Thus, it is ex-ante difficult to predict whether a business hires workers. Related work by de Mel, McKenzie and Woodruff (2010) finds that owners of businesses that do not hire (i.e., own account workers) more closely resemble wageworkers than owners of businesses that hire workers (i.e., employers). Their findings suggest a positive selection of people into wage work as opposed to own-account work.[40]

Importantly, the longitudinal analysis suggests that transitions from own-account worker to employer are limited. We focus on a sample of businesses that initially do not hire workers. Conditional on not hiring initially, we then create an indicator that takes the value one for business owners that start hiring workers, and zero otherwise. In particular, we estimate a linear probability model of the indicator that a manager starts hiring workers as a function of the manager's education, gender, age, urban versus rural residence, and ethnic minority status. We also include survey year and province fixed effects, so that comparisons are made among owners of businesses within the same local market. Table 5 presents the results for the overall sample, as well as rural

[39] This is higher than in the business analysis since owners that are managing a business as their secondary job are not the focus of the current analysis. Businesses that are a secondary job are more likely to be small and less likely to hire workers, which is why the estimate of the share of businesses hiring is greater than in this analysis.
[40] McCaig and Pavcnik (2015) show that there is positive selection of workers from informal to formal work.

and urban areas. Because the results are similar for urban and rural areas, we focus the discussion on results based on overall pooled sample in column 1.

Transitioning to a hiring business is rare: among the owners not initially hiring, only 5.8% start hiring workers. Moreover, the regression analysis in Table 5 suggests that owner characteristics are strongly correlated with the probability of transitioning to a hiring business. Better educated, male, middle age (30-39 year olds), urban, and ethnic minorities are more likely to transition to becoming an owner that hires workers. However, as is the case in the cross sectional analysis presented in Table 4, most of the heterogeneity in transitions to becoming a hiring owner is unexplained. Thus, ex-ante is it difficult to predict which owners will successfully transition and start hiring workers.

The above analysis suggests that transitioning from a non-hiring to a hiring owner (i.e., from own-account owner to employer) is rare. However, some owners might hire workers from the beginning of starting a business. Our data includes all panel households and their members—those that currently have a business and those that do not. As a result, we can also examine the incidence of starting a hiring business and how characteristics of a potential business owner affect the probability that they form a business that immediately hires workers. In our analysis, we use a sample of individuals age 20 to 64 who did not manage a business as their primary job at the start of the two-survey panel. Several interesting findings emerge. First, starting a hiring business is extremely rare: less than 1% of individuals start a business that immediately hires paid workers. In addition, individual characteristics are strongly correlated with the probability that an individual starts a hiring business. We use a linear probability model (see Table 6) to examine how owner characteristics affect probability that an owner starts a business that immediately hires paid works. We regress an indicator that a manager starts a business that immediately hires paid workers on the usual manager's characteristics discussed previously, as well as survey and province fixed effects, and indicators for the previous occupation of the business owner. We include these to examine whether past occupation relates to the ability of an individual to start a hiring business. The past occupational groups include wageworker in agriculture, self-employed outside of agriculture not manager, wageworker in the household business sector, wageworker in the enterprise sector, not in the workforce, with self-employment in agriculture as the excluded category. Past employment is a strong correlate of starting a hiring business. Relative to individuals that were previously self-employed in agriculture, wageworkers in the household sector or self-

employed individuals (but not a manager/owner) in non-agriculture are more likely to start a hiring business. However, wageworkers in the enterprise sector are less likely to start a hiring business than those self-employed in agriculture. This likely reflects positive selection of individuals into wage work in the enterprise sector. Note that past occupation is an important predictor even though we control for education of the individual. In addition, we find that better educated, male, middle age (30-39 year olds), urban residents, and non-ethnic minorities are more likely to directly start a business that hires workers.

Overall, the analysis above suggests that the owner's characteristics are an important correlate of a business' ability to transition from a non-hiring to hiring business or to directly start a hiring business, with the more educated, middle-age, male, non-ethnic minority individuals living in urban areas facing a greater probability of successfully expanding their businesses.

6. Conclusion

We use nationally representative longitudinal household surveys that match businesses to their owners to examine the dynamics of non-farm businesses owned by households in a low-income country. These businesses employ a large share of the workforce in low-income countries. In Vietnam, they employ 60% of the non-agricultural workforce. Consequently, learning about the dynamics of these businesses is key to understanding generation of new jobs and aggregate labor productivity in a low-income country setting.

We find that older businesses do not perform substantially better than their younger counterparts do, which is consistent with patterns observed among formal firms in countries such as India and Mexico (Hsieh and Klenow (2011)). The lack of firm growth with age reflects two underlying forces. First, the flat age profile of businesses is not due to lack of exit of less efficient businesses as exit is prevalent and inversely related to business performance. However, the entering firms do not appear to perform better than the exiting firms do. Consequently, poorly performing exiting firms are replaced by equally poorly performing entrants, which have little hope of growing into a successful business. Second, initially more successful household businesses display or eventually adapt more of the characteristics associated with more formal interactions with government officials and labor markets and characteristics associated with firms in the formal enterprise sector. They are more likely to start hiring workers, increase their workforce and some of them even register as a formal enterprise. However, businesses that undertake these activities are

rare. For example, only about 10% of firms hire workers, and of the incumbents that do not hire workers, less than 5% hire workers in the subsequent period. Transitions of household businesses to formal private enterprises are even rarer. Less than 3% of businesses are registered as an enterprise and among the incumbents initially not registered as an enterprise, only 1.7% of them make this transition in the subsequent period. Our evidence, based on a nationally representative sample, is thus consistent with evidence from several randomized control trials (de Mel, McKenzie, and Woodruff (2010), Schoar (2010)).

Our analysis also highlights important differences in characteristics of a key production input, the owner/manager, across businesses, and these characteristics strongly correlate with business performance. Business growth is positively related to owner's education, but individuals with low levels of education run many of these businesses. The most educated individuals are not business owners, but wageworkers in the formal enterprise sector. Household businesses started by owners who were previously wageworkers in the enterprise sector are relatively rare. Most owners entering or exiting the household business sector are doing so from or to self-employment in agriculture and have low levels of education, especially relative to wageworkers in the formal enterprise sector. Business owners have levels of education comparable to wage workers in the informal sector. Second, consistent with de Mel, McKenzie, and Woodruff (2010), we find that owners that hire workers (i.e., employers) tend to be much better educated than owners that do not hire (own-account owners). Large education gaps suggests that it might be difficult for a non-hiring business to transition to a hiring one, potentially explaining the lack of transitions to hiring businesses discussed above. In fact, the longitudinal dimension of the nationally representative data highlights that transitions from own-account worker to employer are very rare as less than 6% of incumbent businesses that do not initially hire workers do so in the subsequent period. These transitions are more likely to occur in businesses run by better educated owners. Likewise, better-educated owners are more likely to start a business that starts as a hiring business. Finally, even owners that run a business that hires workers (employers) have substantially lower levels of education than wageworkers in the formal enterprise sector. Overall, this analysis emphasizes the strong correlation of an owner's education with business performance and highlights the relatively low levels of education among non-farm business owners, especially relative to the wage earners in the formal sector.

For most businesses owned by households, there is little hope of growth and employment expansion. Expanding businesses that started informal and transition to the formal sector and hire workers are rare, even in countries such as Vietnam, which experienced fast aggregate output growth. This has important consequences for understanding the dynamics of the aggregate share of informal employment in low-income countries. During the past decade, Vietnam experienced a large reduction in the share of employment in informal firms (McCaig and Pavcnik (2015)). Our analysis suggests that this relative decline is not driven by formalization of previously informal household businesses. Instead, it is happening either through the expansion of existing formal firms in the enterprise sector and/or direct entry of firms to the enterprise sector. Understanding the drivers of expansion of the formal sector might thus be a fruitful future area of research. Recent work emphasizes the importance of international trade and deregulation of small-scale manufacturing in this process (McCaig and Pavcnik (2014), Martin, Nataraj, and Harrison (2014)).

Finally, our analysis highlights the usefulness of using nationally representative multi-purpose household surveys with special household business modules to study the dynamics of household businesses. Data constraints often preclude analysis of household business dynamics in low-income countries (Pages (2012), Woodruff (2012)). While our work focuses on Vietnam, several other livings standard measurement surveys from low-income countries contain such modules and provide useful data to explore the dynamics of household businesses in other low-income contexts in future work.

References

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. (2013). "Boom, busts, and household enterprise: Evidence from coffee farmers in Tanzania," mimeo.

Banerjee, Abhijit V. (2013). "Microcredit under the microscope: What have we learned in the past two decades, and what do we need to know?" *Annual Review of Economics*, 5, 487-519.

Banerjee, Abhijit V., and Esther Duflo. (2007). "The Economic Lives of the Poor," *Journal of Economic Perspectives*, 21(1): 141-168.

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman. (2015). "Six randomized evaluations of microcredit: Introduction and further steps," *American Economic Journal: Applied Economics*, 7(1), 1-21.

Bruhn, Miriam and David McKenzie. (2014). "Entry regulation and the formalization of microenterprises in developing countries," *The World Bank Research Observer*, 29(2), 186-201.

Cling, Jean-Pierre, Mireille Razafindrakoto, and François Roubaud (2010). "The informal economy in Vietnam."

Cling, Jean-Pierre, Mireille Razafindrakoto, and François Roubaud (2012). "To be or not to be registered? Explanatory factors behind formalizing non-farm household businesses in Vietnam," *Journal of the Asia Pacific Economy*, 17(4), 632-652.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. (2010). "Who Are the Microenterprise Owners? Evidence from Sri Lanka on Tokman versus De Soto," In Josh Lerner and Antoinette Schoar (Eds.), *International Differences in Entrepreneurship* (63-87). Chicago: University of Chicago Press.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. (2013). "The demand for, and consequences of, formalization among informal firms in Sri Lanka," *American Economic Journal" Applied Economics*, 5(2), 122-150.

de Mel, Suresh, David McKenzie, and Christopher Woodruff (2014). "What Generates Growth in Microenterprises? Experimental Evidence on Capital, Labor and Training," CAGE Working Paper Series No 212.

Goldberg, Pinelopi K., and Nina Pavcnik. (2004). "Trade, inequality, and poverty: What do we know? Evidence from recent trade liberalization episodes in developing countries," *Brookings Trade Forum: 2004*, 223-269.

Goldberg, Pinelopi K., and Nina Pavcnik. (2007). "Distributional Effects of Globalization in Developing Countries," *Journal of Economic Literature*, 45(1), 39-82.

Gollin, Douglas. (2008). "Nobody's business but my own: Self-employment and small enterprise in economic development," *Journal of Monetary Economics*, 55(2), 219-233.

Hsieh, Chang-Tai and Peter J. Klenow. (2009). "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124(4), 1403-1448.

Hsieh, Chang-Tai, and Peter Klenow (2011). "The Life Cycle of Plants in India and Mexico," *The Quarterly Journal of Economics* 129, no.3.

Hsieh, Chang-Tai, and Benjamin A. Olken. (2014). "The Missing "Missing Middle"," *Journal of Economic Perspectives*, 28(3), 89-108.

Kaboski, Joseph P., and Robert M. Townsend. (2011). "A structural evaluation of a large-scale quasi-experimental microfinance initiative," *Econometrica*, 79(5), 1357-1406.

Kaboski, Joseph P., and Robert M. Townsend. (2012). "The impacts of credit on village economies," *American Economic Journal: Applied Economics*, 4(2), 98-133.

La Porta, Rafael, and Andrei Shleifer. (2008). "The Unofficial Economy and Economic Development," Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, 39(2), pages 275-363.

La Porta, Rafael, and Andrei Shleifer. (2014). "Informality and Development," *Journal of Economic Perspectives*, 28(3), 109-26.

Malesky, Edmund and Markus Taussig (2009). "Out of the gray: The impact of provincial institutions on business formalization in Vietnam," *Journal of East Asian Studies*, 9(2), 249-290.

Maloney, William F. (2004). "Informality Revisited," *World Development*, 32(7), 1159-78.

Martin, Leslie A., Shanthi Nataraj, and Ann E. Harrison. (2014). "In with the Big, Out with the Small: Removing Small-Scale Reservations in India," NBER Working Paper No. 19942.

McCaig, Brian, and Nina Pavcnik. (2014). "Export markets and labor allocation in a low-income country," NBER Working Paper No. 20455.

McCaig, Brian and Nina Pavcnik. (2015). "Informal Employment in a Growing and Globalizing Low-Income Country," *American Economic Review*, 105(5), 545-50.

McKenzie, David and Christopher Woodruff. (2015). "Business practices in small firms in developing countries," NBER Working Paper No. 21505.

Melitz, Marc J. (2003). "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 71(6), 1695-1725.

Meltiz, Marc J. and Stephen J. Redding. (2014). "Heterogeneous firms and trade," Handbook of International Economics, 4th Edition, 4:1-54, Elsevier.

Mondragón-Vélez, Camilo, and Ximena Peña (2010). "Business Ownership and Self-Employment in Developing Economies: The Colombian Case," In Josh Lerner and Antoinette Schoar (Eds.), *International Differences in Entrepreneurship* (89-127). Chicago: University of Chicago Press.

Nataraj, Shanthi. (2011). "The impact of trade liberalization on productivity: Evidence from India's formal and informal manufacturing sectors," *Journal of International Economics*, 85(2), 292-301.

Pages, Carmen. (2012). "What are the most important research questions on the economics of growth and labor markets in low-income countries?" First GLM | LIC workshop: Setting the Research Agenda.

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

Paulson, Anna L., Robert Townsend, and Alexander Karaivanov. (2006). "Distinguishing limited liability from moral hazard in a model of entrepreneurship," *Journal of Political Economy*, 114(1), 100-144.

Pavcnik, Nina. (2002). "Trade liberalization, exit, and productivity improvement: Evidence from Chilean plants," *Review of Economic Studies*, 69(1), 245-276.

Porta, Rafael, and Andrei Shleifer (2014). "Informality and Development," *Journal of Economic Perspectives*, 28(3), 109-26.

Restuccia, Diego, and Richard Rogerson. (2008). "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11(4), 707-720.

Samphantharak, Krislert and Robert M. Townsend. (2012). "Measuring the return on household enterprise: What matters most for whom?" *Journal of Development Economics*, 98(1), 58-70.

Schoar, Antoinette. (2010). "The Divide between Subsistence and Transformational Entrepreneurship," In Lerner, Josh and Scott Stern (Eds.), *Innovation Policy and the Economy*, (pp. 57-81). Chicago: University of Chicago Press.

Taussig, Markus and … (2004). "Private enterprise formality and the role of local government," ABD Making Markets Work Better for the Poor Discussion Paper No. 2.

Vijverberg, Wim, Hoang, T. T. H., Nguyen, C.T., Nguyen, N.Q., Nguyen, T.Q., Phung, D.T., and Vu, T.K.M (2006). "Non-farm household enterprises in Vietnam: A research project using data from VHLSS 2004, VHLSS 2002 and AHBS 2003," Research report. Hanoi: The World Bank.

Woodruff, Christopher (2012). "The other half: What we do(n't) know about self-employment in LICs?" First GLM|LIC workshop: Setting the research agenda.

Figure 1: Heterogeneity of household business performance

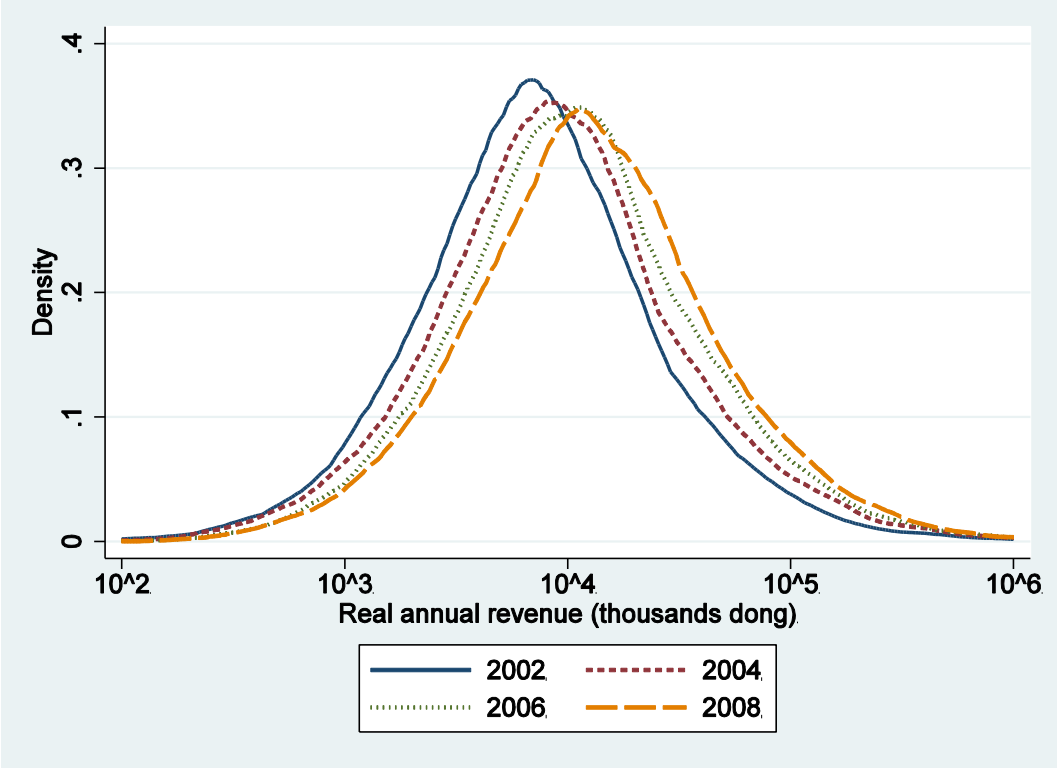


Figure 2: Age of household businesses

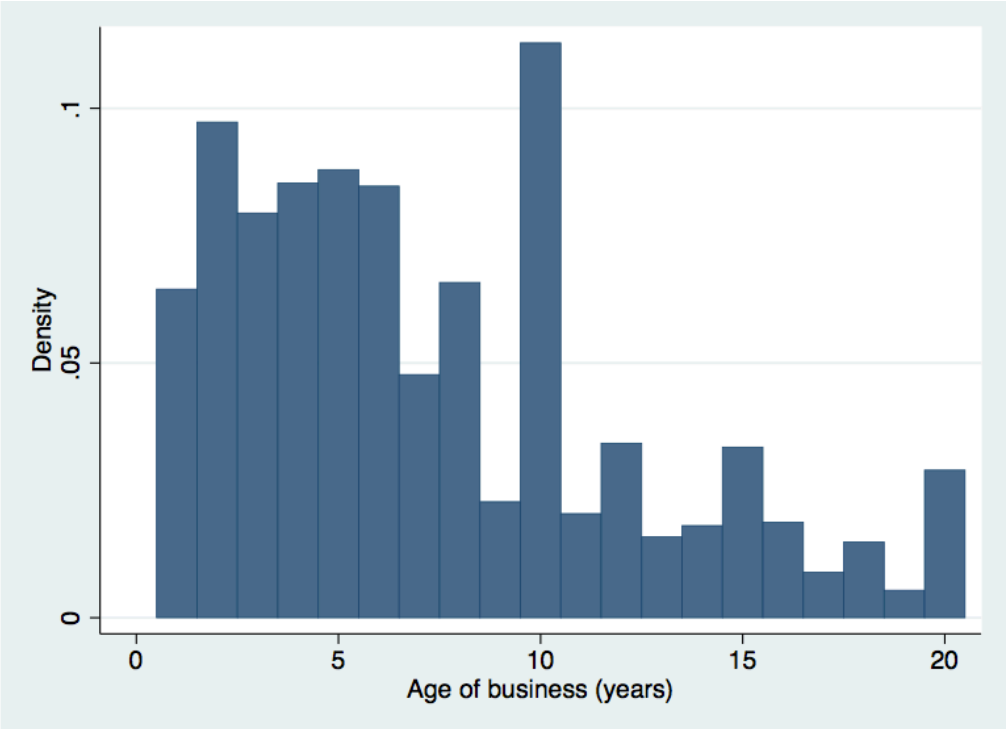


Figure 3a: Older businesses do not employ more workers

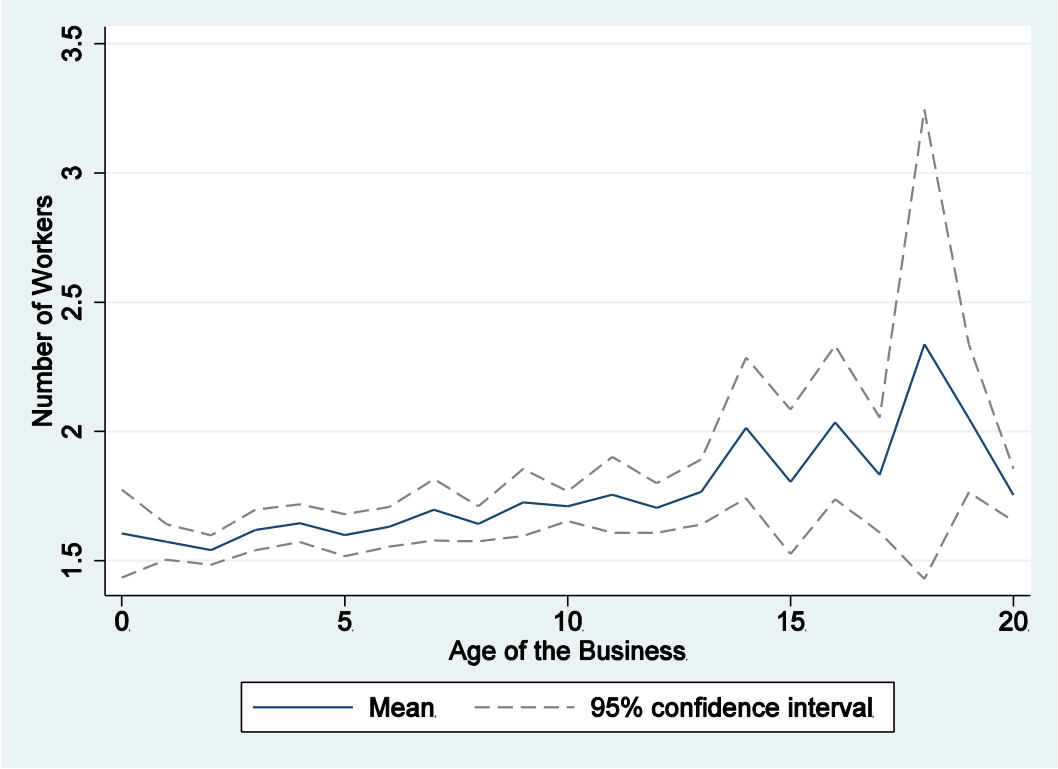


Figure 3b: Older businesses do not have higher revenue

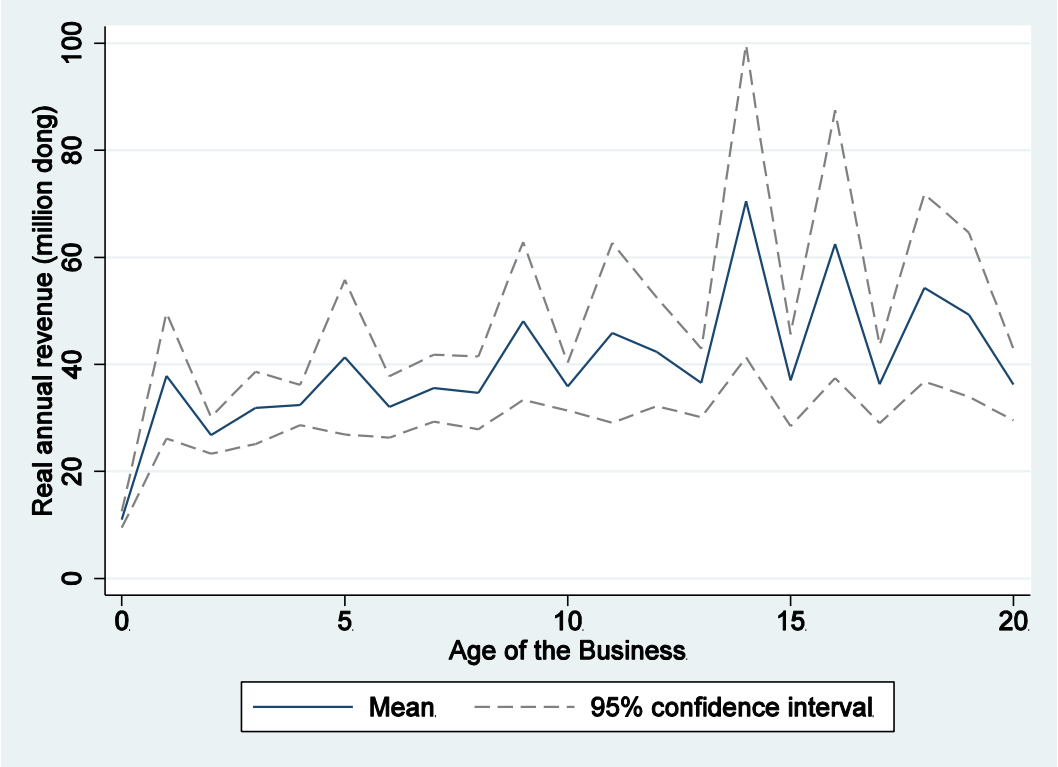


Figure 3c: Older businesses do not have higher revenue per worker

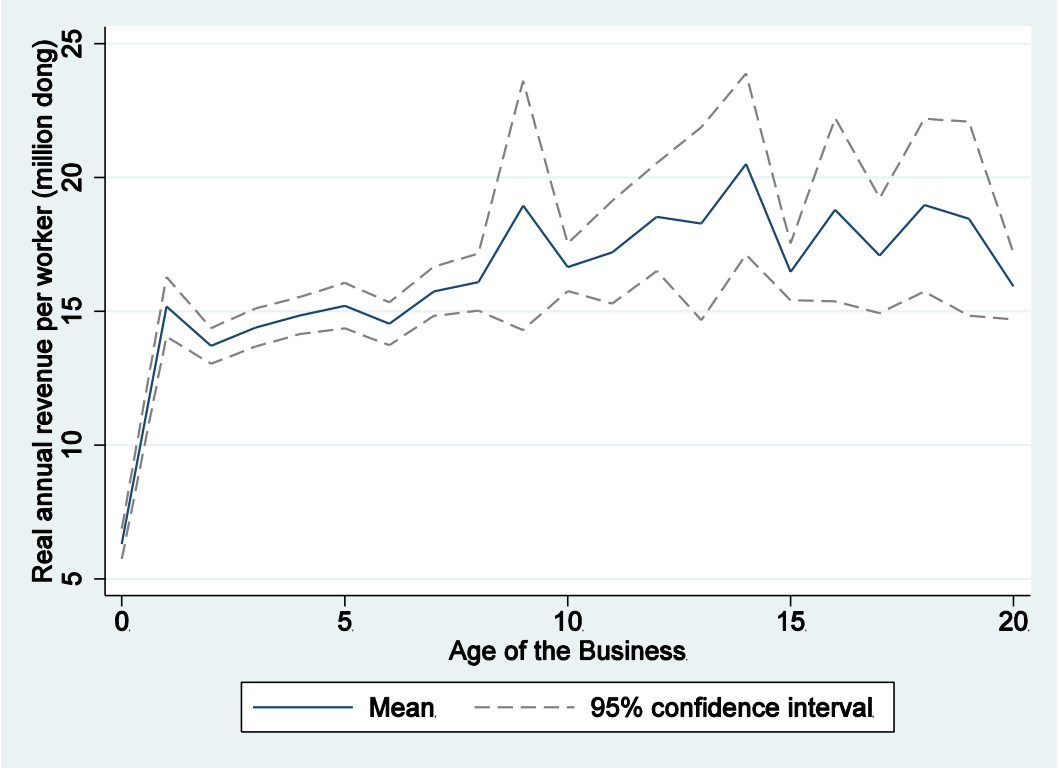
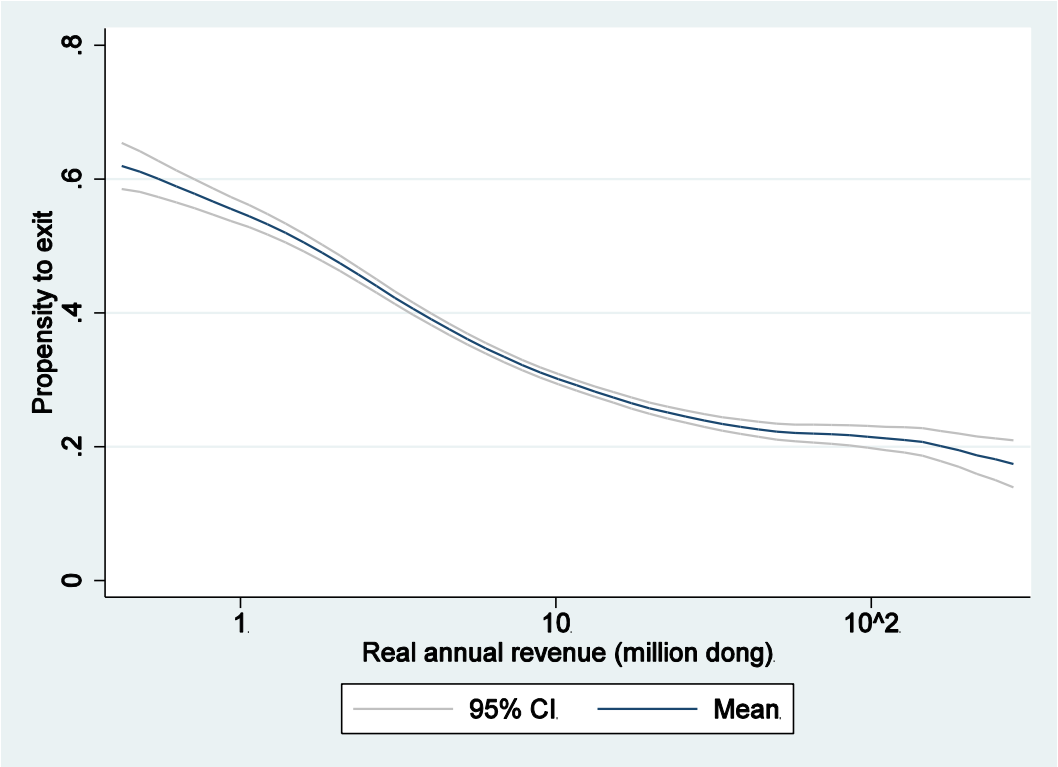


Figure 4: Selection at work: less successful household businesses more likely to exit



Note: The bottom and top 1% of observations based on revenue have been trimmed in the figure.

Figure 5a: Business performance: incumbents, entrants, and exiters

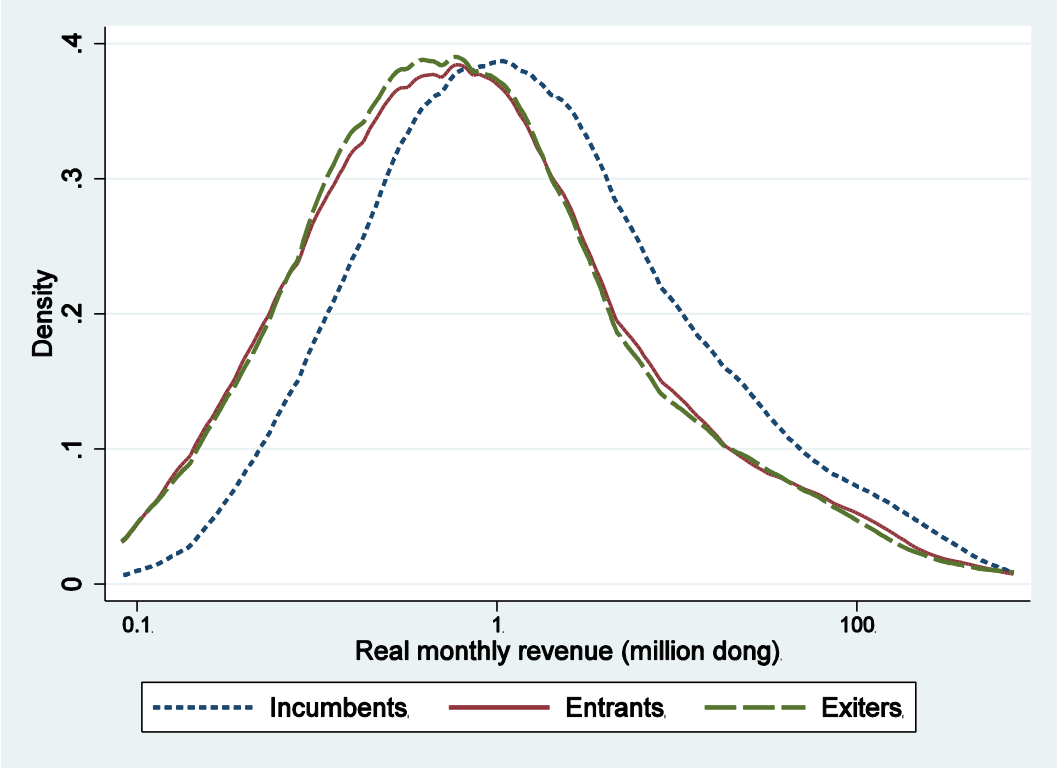
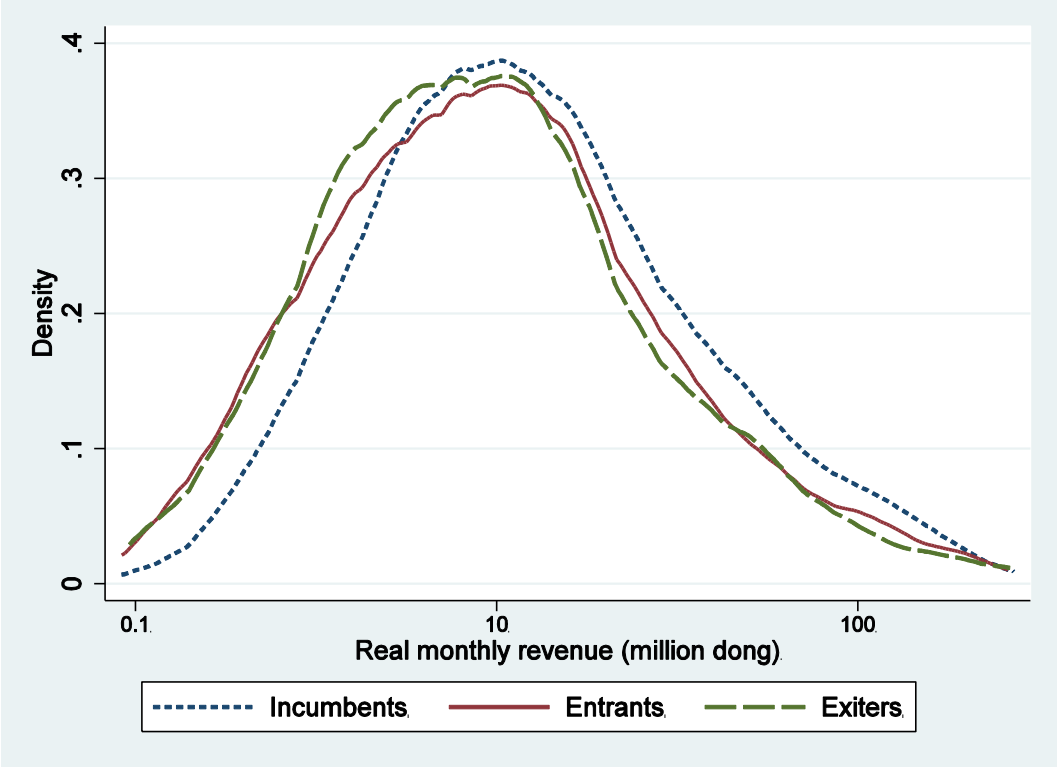
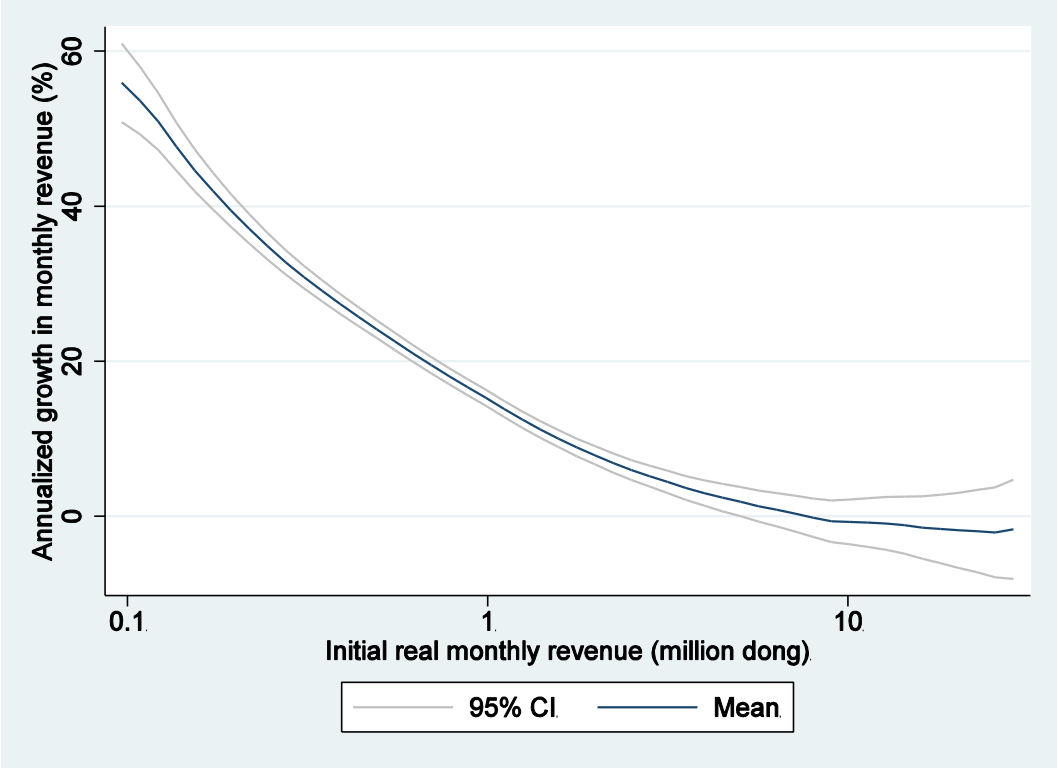


Figure 5b: Business performance: Incumbents, entrants, and exiters



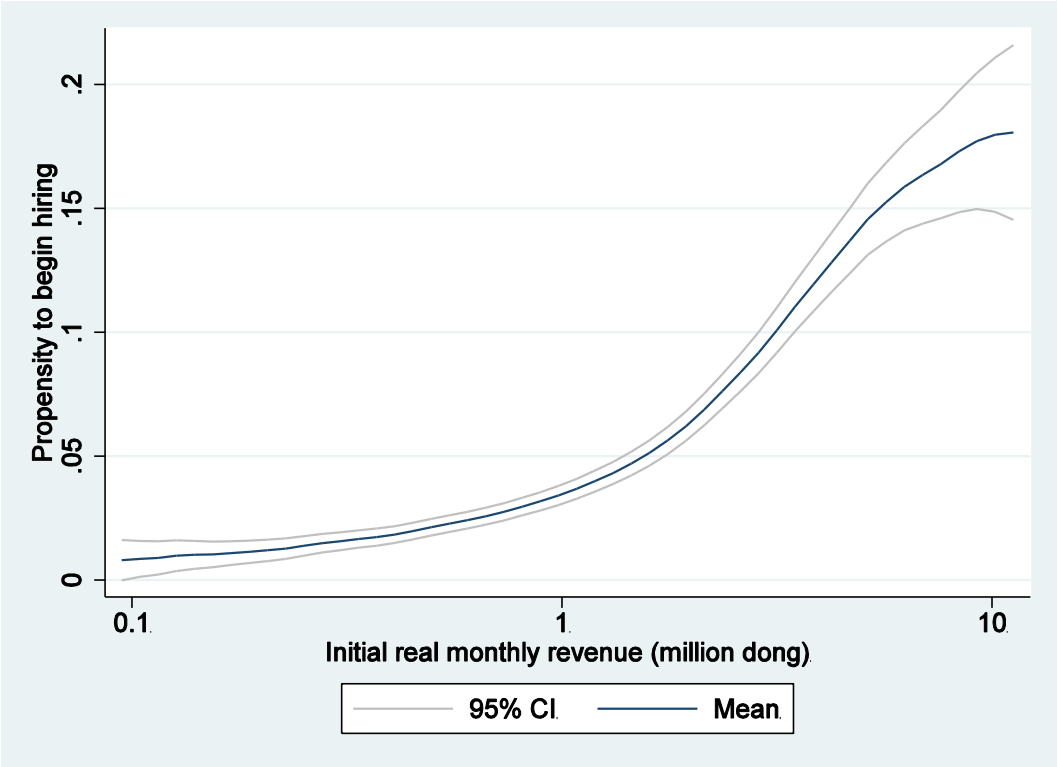
Note: This graph excludes businesses that entered and exited during the period.

Figure 6: Initial performance and subsequent revenue growth



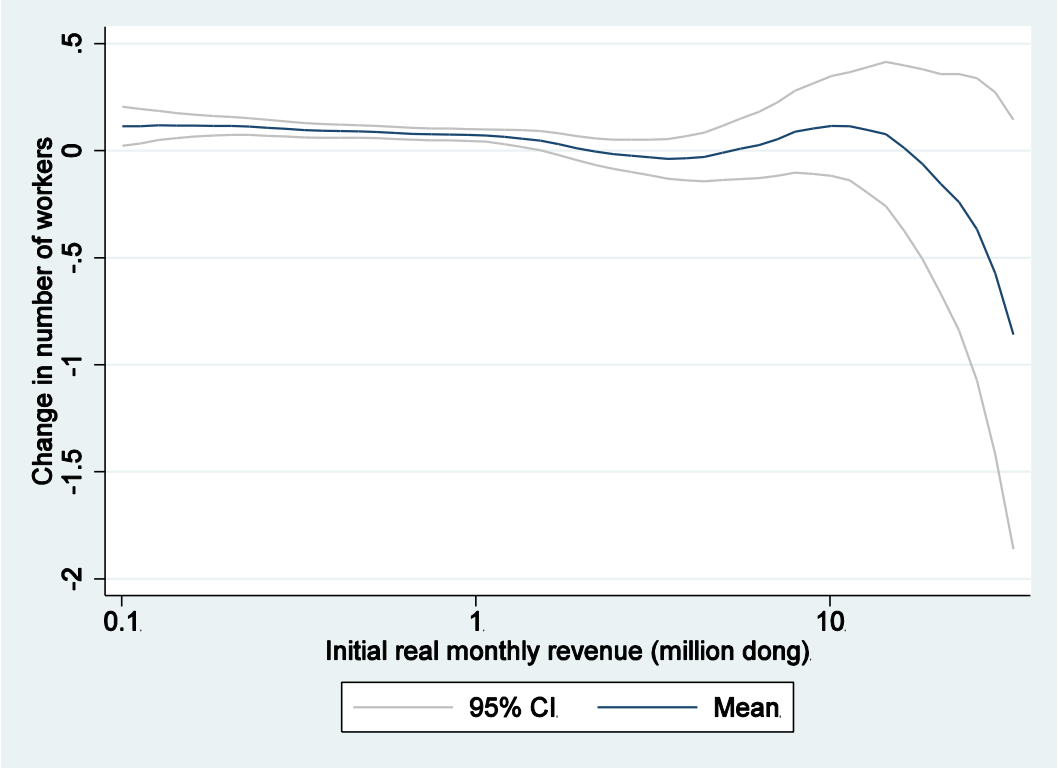
Note: The bottom and top 1% of observations based on revenue have been trimmed in the figure.

Figure 7a: Selection among incumbents: Hiring outside labor



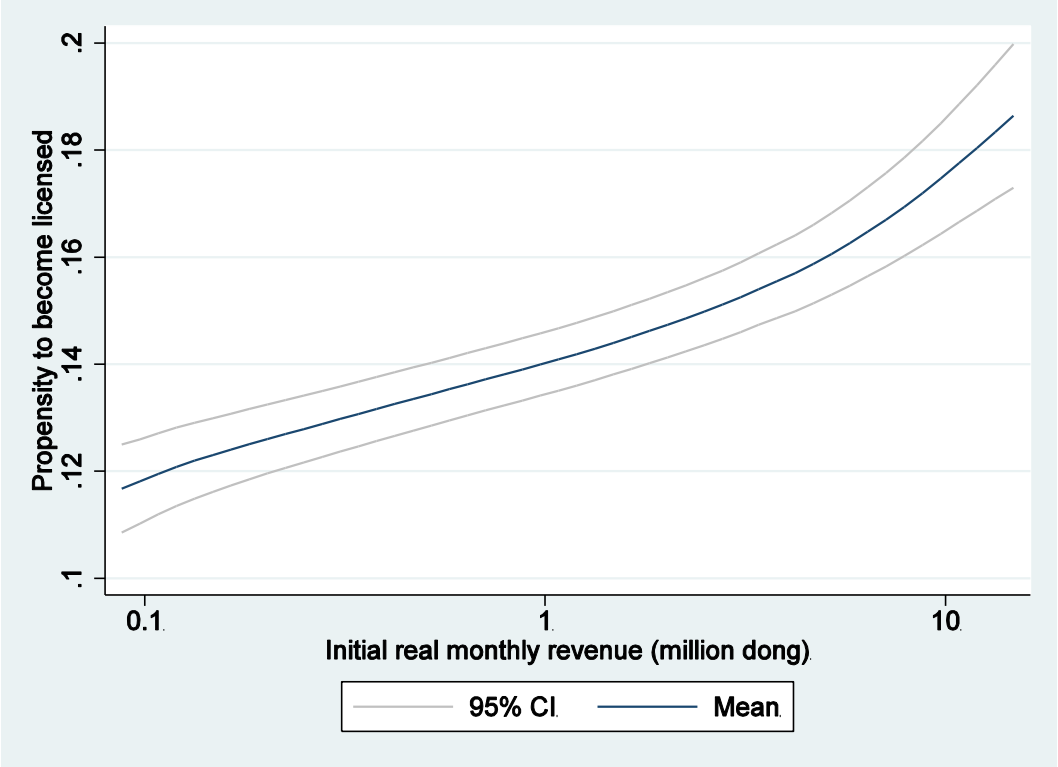
Note: The bottom and top 1% of observations based on revenue have been trimmed in the figure.

Figure 7b: Selection among incumbents: Change in the number of workers



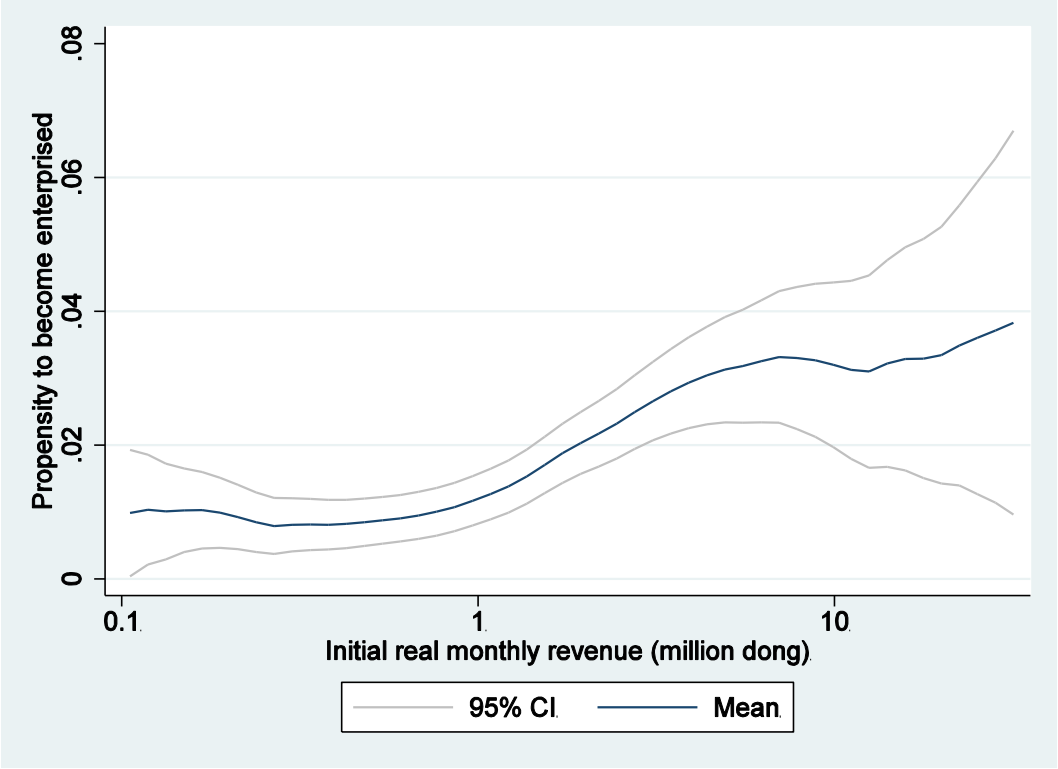
Note: The bottom and top 1% of observations based on revenue have been trimmed in the figure.

Figure 8a: Selection among incumbents: obtaining a business license



Note: The bottom and top 1% of observations based on revenue have been trimmed in the figure.

Figure 8b: Selection among incumbents: formalization



Note: The bottom and top 1% of observations based on revenue have been trimmed in the figure.

Figure 9a: Owner's education and change in the number of workers

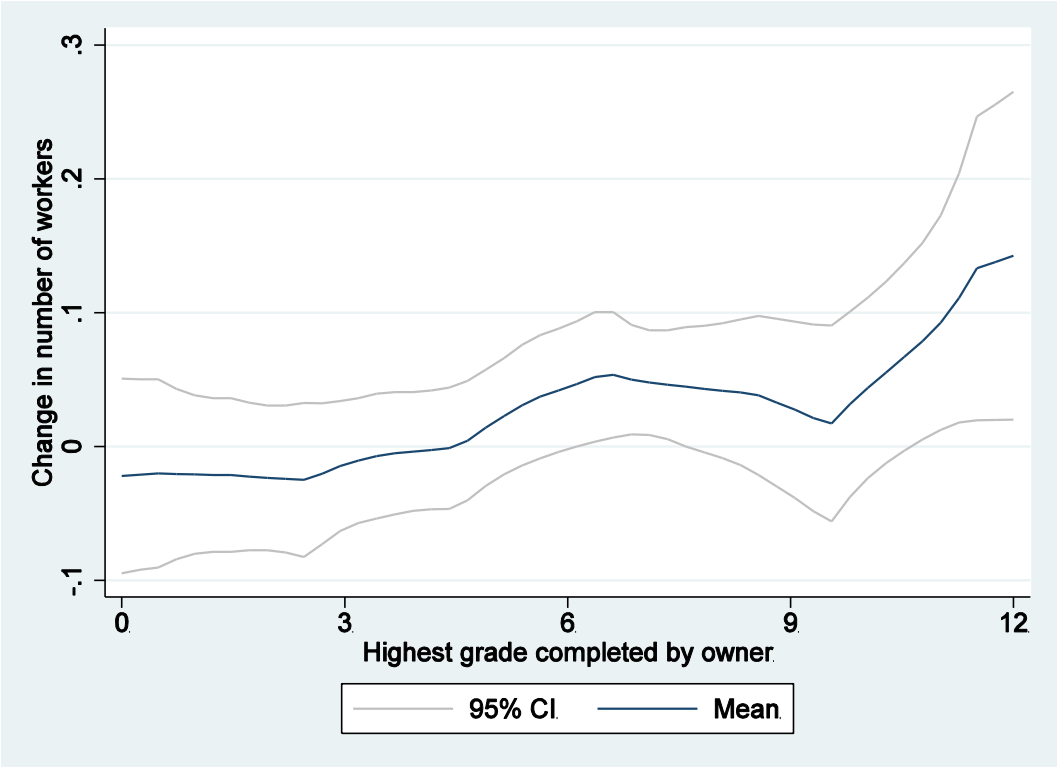
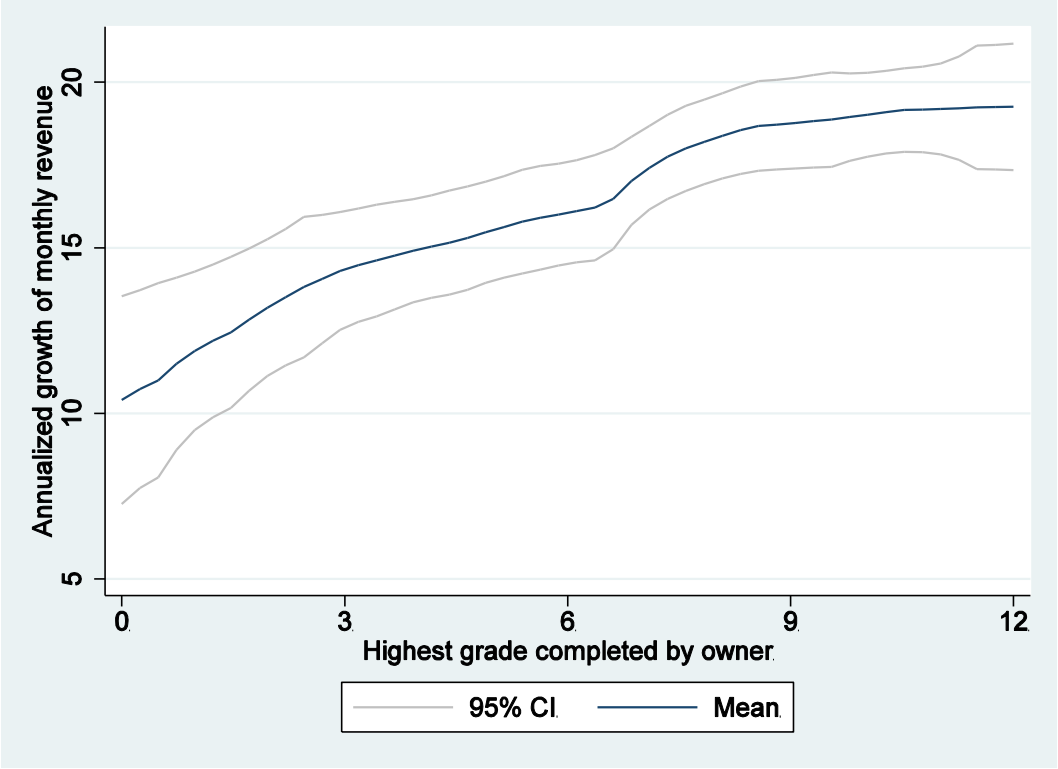


Figure 9b: Owner's education and revenue growth



Data Appendix A

In this appendix, we describe in detail three data preparation steps. We first describe how we verify that the correct individual within the household has been reported as the manager for the respective business. Second, we describe how we predict a manager for all businesses in the 2004 VHLSS. Unlike the 2002, 2006, and 2008 VHLSS, respondents in 2004 were not asked to identify the most knowledgeable individual for the business (hereafter referred to as the manager for brevity). Third, we describe how we create a household business panel.

A.1. Verifying information about the manager of a household business

The primary activity has been verifying whether the correct individual within the household has been reported as the manager for the respective business. We checked whether the manager reports information in the employment module that is consistent with managing the household business reported in the household business module of the survey. In particular, the manager should report working, should report working in a household business, and report working in the same industry as the business if sufficient detail is provided for the job. In the 2002 VHLSS, detailed information is provided for the primary job only, whereas in the 2004, 2006, and 2008 VHLSSs detailed information is also provided for the secondary job. As an additional step, we also checked for instances in which an individual reported working in a household business in the employment module but there was no business reported by the household with a matching industry.

Across the three surveys for which the manager was reported in the business module, we find that the vast majority of reported managers provided consistent information in the labor module. Thus, we have changed the manager for only 5.2, 2.2, and 2.4 percent of businesses in 2002, 2006, and 2008 respectively. We find it very reassuring that the vast majority of reported managers provided consistent data in the labor module.

A.2. Predicting the manager of a household business in 2004 VHLSS

The 2004 VHLSS household business module does not include information on the most knowledgeable person/business manager as the 2002, 2006, or 2008 VHLSSs. Hence, we predicted the manager of each business in 2004.

This is useful for three reasons. First, part of our analysis requires looking at the relationship

between various business performance indicators and characteristics of the manager, such as age, education, prior work experience, etc. Second, knowing the manager of the business helps to facilitate the matching of businesses over time. Third, knowing the manager of the business in 2004, we can use information from the employment module to identify several other characteristics of a businesses in 2004. For example, we can determine whether a business is a household business or a private enterprise. In the employment module, the individual reports whether they were self-employed in an enterprise or self-employed in a household business, which enables us to make a distinction between a household business and a private enterprise in the 2004 data. Moreover, in the employment module, the individual also reports how long they have been doing the job and thus we can infer a possible year of start for the business.

We have developed an algorithm, which has a very high rate of success, 92.9%, when tested using the 2006 VHLSS. Below we provide a detailed description of the data available in the employment modules and business modules of the 2004 VHLSS which can be used for matching individuals with businesses, the algorithm used for matching, a summary of how the matches were made, and the percentage of successful predictions from using the same algorithm on data from the 2006 VHLSS.[41]

We combine data from the employment and business modules of the 2004 VHLSS, which can be matched. In particular, from the employment module we identify individuals that reported being self-employed in a household business for either their primary or secondary job during the past year. For these jobs, we use information on the industry, the number of months worked during the past years, the number of days per month usually worked, and the number of years the individual has been doing the job. From the business module, we use information on the industry, the number of months operating during the past year, the average number of days per month operating, and the year the business started.[42]

In Table A1, we provide a summary of the matches by the step within the manager prediction algorithm at which the match was made. The table is organized sequentially such that the first step of the algorithm was to identify the manager for businesses in which only one

[41] We did not use the 2002 VHLSS for testing the algorithm because it did not collect as much information about an individual's secondary job as the 2004 and 2006 VHLSSs did. Since many businesses are run as a second job, testing the algorithm using the 2006 VHLSS is more appropriate.

[42] The year the business started is only available for about 1/5th of the sample since this question was not asked of all businesses, but instead was part of an extra module on businesses that only 1/5th of households were asked.

household member reported being self-employed in the industry of the business and then only businesses without a predicted manager would proceed to the next row. The first step of the algorithm matches an individual as the manager for the business for 70.5% of all businesses in the 2004 VHLSS. The corresponding rate of success using the 2006 VHLSS is 98.9%. Thus, for a large share of businesses we have a very high degree of confidence in our predicted manager. Next, we identified a manager for any remaining businesses when there was only one household member that reported being a manager of a business in the same industry in the 2006 VHLSS and so on down the rows of the table.[43] In sum, the algorithm correctly identified the manager for 92.9% of businesses in the 2006 VHLSS. Thus, our manager prediction algorithm is doing a very good job of identifying the manager of the business.[44]

A.3 Creation of Panel of Businesses

In this section, we explain how we match businesses in the two- and three-survey VHLSSs panels. The surveys are household-level panels. The household surveys were not designed to directly follow businesses and thus we use characteristics of the business that should not change for most businesses in order to match them over time. We use the longitudinal dimension of our data at the household and individual level.

Note that not all businesses run by a panel household should be matched over time. For example, any household that reports running a different number of businesses across the two years has experienced net entry or exit of businesses and thus at least one business within the household should not be matched. Thus, for any given household the maximum number of matched businesses is the minimum of the number of businesses run in either year. Table A2 summarizes the number of businesses run by panel households in each of the two-survey panels. There are 65,134

[43] Note that the percentage of successfully identified managers in the 2006 VHLSS for "Only household member with matching months and days per month" is likely an underestimate of the rate for the 2004 VHLSS. This is because only about 1/5th of businesses in the 2004 VHLSS have information on the year when the business started whereas all businesses in the 2006 VHLSS have this information. Thus, some 2004 businesses for which the year was not reported, but the number of years, months, and days all matched would only be matched in the row "Only household member with matching months and days per month". Indeed, in the 2006 VHLSS 7.3% of businesses are matched in the step "Only household member with matching years, months, and days per month" as compared to only 1.1% in the 2004 VHLSS and 2.0% of 2006 VHLSS businesses were matched to a manager in the step "Only household member with matching months and days per month" as compared to 5.8% in the 2004 VHLSS.

[44] Our algorithm does not predict a manager for 131 out of 21,458 (0.6 percent) businesses. This could be due to the business being managed by an individual as their third job, which our algorithm currently does not include, or due to measurement error in either the industry of the business or the industry of the job.

panel households in our dataset. A little over half of the households did not operate a business in either the start or end year of the respective panel. The number of businesses that can potentially be matched is 21,556.[45]

We start by matching businesses using information on the industry of operation and the manager of the business. We match 14,229 businesses based on these matching criteria or 66 percent of the maximum possible number of matches. We subsequently relax the matching criteria and consider matching the remaining unmatched businesses by either industry (allowing the manager of the business to change over time) or by manager (allowing the industry of the business to change over time). Matching by manager leads to an additional 4,045 matches (additional 18.8 percent) while matching by industry leads to additional 1,660 matches (additional 7.7 percent).

[45] This is derived by summing over $\min(i,j)*a_{ij}$ where i represents the number of businesses run by the household in the start year of the panel, j is the number of businesses run by the household in the end year of the panel, and a_{ij} is the number of households operating i businesses in the start year and j businesses in the end year.

Table 1a: Number of households operating businesses

	2002	2004	2006	2008
Total number of households	74,350	45,928	45,943	45,945
Number of households …				
Operating businesses	27,812	17,293	16,564	16,559
Operating 1 business	21,750	13,684	13,152	13,134
Operating 2 businesses	5,295	3,120	2,980	3,000
Operating 3 businesses	665	422	382	369
Operating 4 businesses	102	67	50	56

Source: Authors' calculations using the 2002, 2004, 2006, and 2008 Vietnam Household Living Standards Surveys.

Table 1b: Business and manager characteristics summary

	2002 Mean	2002 St. Dev.	2004 Mean	2004 St. Dev.	2006 Mean	2006 St. Dev.	2008 Mean	2008 St. Dev.
Number of businesses	34,743	34,743	21,458	21,458	20,458	20,458	20,465	20,465
Share in primary industries	0.013	0.115	0.008	0.090	0.007	0.086	0.006	0.080
Share in secondary industries	0.282	0.450	0.278	0.448	0.265	0.441	0.250	0.433
Share in tertiary industries	0.704	0.456	0.714	0.452	0.728	0.445	0.744	0.436
Share in urban areas	0.339	0.473	0.331	0.471	0.342	0.475	0.392	0.488
Share with a household business license	0.193	0.394	0.213	0.409	0.240	0.427	0.265	0.442
Share registered as an enterprise					0.023	0.150	0.022	0.148
Mean months of operation	9.75	3.02	10.09	2.89	10.29	2.74	10.48	2.67
Mean real revenue	20,079	106,449	32,122	300,903	37,296	275,071	42,386	216,523
Mean real expenses	10,190	80,454	21,226	294,249	24,479	258,126	27,036	201,238
Mean labour expense ratio	0.033	0.121	0.029	0.115	0.034	0.126	0.037	0.131
Share that report labour expenses	0.113	0.316	0.093	0.290	0.100	0.301	0.110	0.314
Mean number of workers			1.68	2.73	1.69	2.47	1.73	3.47
Mean number of paid workers			0.32	2.11	0.38	2.26	0.41	3.37
Mean age of the business			7.75	6.93	8.12	6.93	8.44	7.04
Share with manager operating business as primary job	0.736	0.441	0.746	0.435	0.763	0.426	0.777	0.417
Share with female manager	0.572	0.495	0.584	0.493	0.584	0.493	0.580	0.494
Mean age of manager	39.3	11.8	40.6	11.7	41.8	11.5	42.9	11.5
Mean grade completed by manager	7.55	3.23	7.69	3.21	7.73	3.24	7.84	3.27
Share with manager with less than primary completed	0.184	0.388	0.172	0.378	0.171	0.376	0.161	0.368
Share with manager with primary but less than lower secondary	0.306	0.461	0.295	0.456	0.291	0.454	0.291	0.454
Share with manager with lower secondary but less than upper secondary	0.344	0.475	0.354	0.478	0.349	0.477	0.343	0.475
Share with manager with upper secondary	0.166	0.372	0.178	0.382	0.188	0.391	0.205	0.404
Share with ethnic minority manager	0.053	0.224	0.054	0.226	0.060	0.237	0.060	0.238

Source: Authors' calculations using the 2002, 2004, 2006, and 2008 Vietnam Household Living Standards Surveys.

Table 2: Mean highest grade completed by occupation

Occupation	Share of workers	Mean grade
Wage worker in agriculture	0.047	5.0
Wage worker in household sector (excluding agriculture)	0.104	7.7
Wage worker in enterprise sector (excluding agriculture)	0.183	10.5
Self-employed in agriculture	0.465	6.6
Self-employed in household business, not manager	0.041	8.1
Self-employed in household business, manager, does not hire	0.139	7.7
Self-employed in household business, manager, does hire	0.020	9.0

Source: Authors' calculations using the 2002, 2004, 2006, and 2008 Vietnam Household Living Standards Surveys. Notes: The sample is all individuals age 20 to 64 that reported working during the past 12 months. The classification of occupation is based on the most time-consuming job during the past 12 months. Estimated means based on sampling weights.

Table 3: Occupations before and after operating a business

Occupation	Before entering	After exiting
Wage worker in agriculture	3.8	3.0
Wage worker in household sector	15.3	16.2
Wage worker in enterprise sector	8.4	9.8
Self-employed in agriculture	41.7	36.9
Self-employed in household business, not manager	17.3	16.4
Not in the workforce	13.5	17.8

Source: Authors' calculations using the 2002, 2004, 2006, and 2008 Vietnam Household Living Standards Surveys. Notes: The table shows the distribution of occupations in the previous survey for individuals that begin managing a business and in the subsequent survey for individuals that cease managing a business. In both cases the sample is restricted to individuals age 20 to 64 at the time of managing the business. Estimated means based on sampling weights from the latter survey.

Table 4: Observeable differences between managers that hire and those that do not
Dependent variable: Indicator for being a manager that hires workers

	(1) Pooled	(2) Rural	(3) Urban
Completed primary but not lower secondary	0.022***	0.020***	0.028***
	(0.004)	(0.005)	(0.007)
Completed lower secondary but not upper secondary	0.052***	0.038***	0.073***
	(0.004)	(0.005)	(0.007)
Completed upper secondary	0.109***	0.063***	0.151***
	(0.005)	(0.006)	(0.008)
Female	-0.099***	-0.112***	-0.084***
	(0.003)	(0.003)	(0.004)
Age 25 to 29	0.035***	0.034***	0.033***
	(0.007)	(0.007)	(0.013)
Age 30 to 39	0.058***	0.057***	0.057***
	(0.006)	(0.006)	(0.011)
Age 40 to 49	0.054***	0.050***	0.058***
	(0.006)	(0.007)	(0.011)
Age 50 to 64	0.029***	0.030***	0.027**
	(0.006)	(0.007)	(0.011)
Urban	0.017***		
	(0.003)		
Ethnic minority	-0.036***	-0.033***	-0.052***
	(0.008)	(0.009)	(0.016)
2004	-0.021***	-0.030***	-0.008
	(0.003)	(0.004)	(0.006)
2006	-0.012***	-0.019***	-0.001
	(0.003)	(0.004)	(0.006)
2008	-0.005	-0.017***	0.009*
	(0.003)	(0.004)	(0.006)
Constant	0.090***	0.146***	0.051***
	(0.010)	(0.014)	(0.016)
Observations	66,204	38,416	27,788
R-squared	0.050	0.052	0.053
Province FE	Yes	Yes	Yes

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Notes: The table presents coefficient estimates from a linear probability model of the probability of being a manager that hires relative to a manager that does not hire workers. The sample is restricted to individuals age 20 to 64 and is based on the primary job of the individual.

Table 5: Probability of starting to hire workers

Dependent variable: Indicator for being a manager that starts to hire workers

	(1) Pooled	(2) Rural	(3) Urban
Completed primary but not lower secondary	0.007	0.012	-0.000
	(0.007)	(0.008)	(0.012)
Completed lower secondary but not upper secondary	0.013*	0.006	0.024**
	(0.007)	(0.009)	(0.012)
Completed upper secondary	0.047***	0.019*	0.073***
	(0.008)	(0.011)	(0.013)
Female	-0.043***	-0.054***	-0.033***
	(0.005)	(0.006)	(0.007)
Age 25 to 29	0.014	0.008	0.027
	(0.013)	(0.015)	(0.025)
Age 30 to 39	0.020*	0.027**	0.013
	(0.011)	(0.013)	(0.023)
Age 40 to 49	0.012	0.015	0.011
	(0.012)	(0.013)	(0.023)
Age 50 to 64	0.000	0.006	-0.004
	(0.012)	(0.014)	(0.023)
Urban	0.008		
	(0.005)		
Ethnic minority	0.009	0.006	0.012
	(0.015)	(0.017)	(0.027)
2004-6 panel	0.003	0.010	-0.006
	(0.005)	(0.007)	(0.009)
2006-8 panel	0.012**	0.006	0.020**
	(0.005)	(0.007)	(0.009)
Constant	0.063***	0.077***	0.055*
	(0.018)	(0.026)	(0.030)
Observations	11,488	6,577	4,911
R-squared	0.025	0.034	0.038
Province FE	Yes	Yes	Yes

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Notes: The table presents coefficient estimates from a linear probability model of the probability of being a manager that starts to hire workers relative to a manager that does not start to hire workers. The sample is restricted to individuals age 20 to 64 who managed a household business that did not hire workers as their primary job at the start of the two-survey panel and managed either a hiring or non-hiring business at the end of the two-survey panel.

Table 6: Probability of becoming the manager of a hiring business

Dependent variable: Indicator for becoming the manager of a hiring business

	(1) Pooled	(2) Rural	(3) Urban
Wage worker in agriculture	-0.002**	-0.002**	-0.004
	(0.001)	(0.001)	(0.004)
Self-employed outside of agriculture, not manager	0.034***	0.029***	0.042***
	(0.001)	(0.001)	(0.003)
Wage worker in household sector (excluding agriculture)	0.011***	0.009***	0.017***
	(0.001)	(0.001)	(0.002)
Wage worker in enterprise sector (excluding agriculture)	-0.002***	-0.001	-0.003
	(0.001)	(0.001)	(0.002)
Not in the workforce	0.000	0.000	0.001
	(0.001)	(0.001)	(0.002)
Completed primary but not lower secondary	0.002***	0.002**	0.004*
	(0.001)	(0.001)	(0.002)
Completed lower secondary but not upper secondary	0.004***	0.003***	0.007***
	(0.001)	(0.001)	(0.002)
Completed upper secondary	0.006***	0.005***	0.010***
	(0.001)	(0.001)	(0.003)
Female	-0.004***	-0.004***	-0.005***
	(0.000)	(0.000)	(0.001)
Age 25 to 29	0.005***	0.004***	0.009***
	(0.001)	(0.001)	(0.003)
Age 30 to 39	0.007***	0.005***	0.010***
	(0.001)	(0.001)	(0.002)
Age 40 to 49	0.004***	0.003***	0.008***
	(0.001)	(0.001)	(0.002)
Age 50 to 64	0.004***	0.003***	0.007***
	(0.001)	(0.001)	(0.002)
Urban	0.003***		
	(0.001)		
Ethnic minority	-0.003***	-0.003***	-0.004
	(0.001)	(0.001)	(0.003)
2004-6 panel	-0.002***	-0.001***	-0.004***
	(0.001)	(0.001)	(0.002)
2006-8 panel	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.002)
Constant	0.001	0.003	-0.002
	(0.002)	(0.002)	(0.004)
Observations	117,910	91,315	26,595
R-squared	0.014	0.011	0.019
Province FE	Yes	Yes	Yes

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Notes: The table presents coefficient estimates from a linear probability model of the probability of becoming the manager of a business that hires workers. The sample is restricted to individuals age 20 to 64 who did not manage a household business as their primary job at the start of the two-survey panel. The dependent variable is based on the job status of the primary job at the end of the two-survey panel.

Table A1: Manager prediction results

Match criteria	Number of matches in 2004	Share of matches in 2004	Share of correct matches 2006
No job matched the business	131	0.006	0.000
Businesses matched to a primary or secondary job			
Only job that matched the business by industry	15,122	0.705	0.989
Only manager in same business in subsequent survey	1,595	0.074	0.755
Only job that matched by year, months and days	232	0.011	0.912
Only job that matched by months and days	1,250	0.058	0.794
Only job that matched by months	191	0.009	0.755
Highest number of years in the job	742	0.035	0.789
Highest number of days worked in the past year in the job	180	0.008	0.659
Only one of the head or spouse matched	186	0.009	0.831
Highest number of hours per day in the job	313	0.015	0.681
Highest ranked individual within household	968	0.045	0.703
Only primary job	3	0.000	1.000
Businesses not matched to a primary or secondary job			
Only third job that matched business	421	0.020	0.952
Only manager in same business in subsequent survey	33	0.002	0.667
Highest ranked individual within household	91	0.004	0.443
Total	**21,458**	**1.000**	**0.929**

Table A2: Number of households by number of businesses

Number of businesses run at start of panel	Number of business run at end of panel					
	0	1	2	3	4	Total
			2002-2004 household panel			
0	11,804	1,959	171	6	1	13,941
1	1,944	3,949	648	76	7	6,624
2	192	697	627	92	8	1,616
3	12	52	101	32	6	203
4	1	12	10	6	2	31
Total	13,953	6,669	1,557	212	24	22,415
			2004-2006 household panel			
0	11,857	1,466	126	10	1	13,460
1	1,643	4,222	581	54	7	6,507
2	163	637	647	75	9	1,531
3	13	46	81	42	3	185
4	4	7	10	8	2	31
Total	13,680	6,378	1,445	189	22	21,714
			2006-2008 household panel			
0	11,830	1,428	133	9	0	13,400
1	1,466	3,911	588	57	5	6,027
2	119	541	639	58	16	1,373
3	8	41	80	42	6	177
4	2	5	10	8	3	28
Total	13,425	5,926	1,450	174	30	21,005