

Lay, Jann; Tafese, Tevin

Working Paper

Formalization and productivity: Firm-level evidence from Viet Nam

WIDER Working Paper, No. 2020/164

Provided in Cooperation with:

United Nations University (UNU), World Institute for Development Economics Research (WIDER)

Suggested Citation: Lay, Jann; Tafese, Tevin (2020) : Formalization and productivity: Firm-level evidence from Viet Nam, WIDER Working Paper, No. 2020/164, ISBN 978-92-9256-921-1, The United Nations University World Institute for Development Economics Research (UNU-WIDER), Helsinki, <https://doi.org/10.35188/UNU-WIDER/2020/921-1>

This Version is available at:

<https://hdl.handle.net/10419/229388>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



UNITED NATIONS
UNIVERSITY
UNU-WIDER

WIDER Working Paper 2020/164

Formalization and productivity

Firm-level evidence from Viet Nam

Jann Lay and Tevin Tafese*

December 2020

Abstract: Using a firm-level panel dataset on private small- and medium-sized enterprises (SMEs) in Viet Nam's manufacturing sector, this paper examines productivity dynamics of formal and informal firms. We decompose productivity changes into changes within and between formal and informal firms. We assess the contributions of firm entry and exit as well as informal–formal transitions. Our results show that productivity is considerably lower and misallocation more prevalent in the informal than in the formal sector. Yet, formalizing firms in Viet Nam make an important contribution to aggregate productivity growth among manufacturing SMEs, growing faster than other firms and increasing efficiency. We identify two ‘regimes’ of formalization. Until early 2010, more productive (previously) informal firms formalize. Policy changes and accelerated formalization then alter the characteristics of formalizers, as less productive firms become formal. While this formalization wave depresses average formal total factor productivity growth, the overall productivity effect is positive.

Key words: formalization, productivity, Viet Nam, misallocation

JEL classification: D22, D24, O11, O17

Acknowledgements: We are grateful for the productive and stimulating collaboration and the support of the United Nations University World Institute for Development Economics Research, Helsinki.

*All authors: GIGA Institute of African Affairs, German Institute for Global and Area Studies (GIGA), Hamburg, Germany and University of Goettingen, Germany; corresponding author: tevin.tafese@giga-hamburg.de

This study has been prepared within the UNU-WIDER project [Transforming informal work and livelihoods](#).

Copyright © UNU-WIDER 2020

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-921-1

<https://doi.org/10.35188/UNU-WIDER/2020/921-1>

Typescript prepared by Ayesha Chari.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

The views expressed in this paper are those of the author(s), and do not necessarily reflect the views of the Institute or the United Nations University, nor the programme/project donors.

1 Introduction

The misallocation of resources and the related dispersion of productivity across firms and farms in developing countries implies that countries turn out to be poorer than they could potentially be given their endowments (Hsieh and Klenow 2009; Restuccia and Rogerson 2017). Productivity differences across firms may often be caused by market imperfections, including on credit and labour markets and informational frictions, and by distortionary policies, such as erratic regulatory practices and exposure to corruption. These constraints are likely to be more binding for informal firms, which are not registered with the authorities, do not pay taxes, and are only partially subject to regulation. In particular, informal firms lack access to contractual enforcement mechanisms and access to formal finance. By putting a break on firm growth informality can cause higher productivity dispersion and thus can negatively affect aggregate growth. Accordingly, formalization (policies) may be able to remove growth constraints to formerly informal firms and improve allocative efficiency. There is a wealth of empirical knowledge on the characteristics and heterogeneity of the informal sector that consists of subsistence activities as well as potentially dynamic but constrained firms (e.g., de Mel et al. 2008; Grimm et al. 2012; La Porta and Shleifer 2014). Today, the majority view on the informal sector is that informality (or formality) is the outcome of a deliberate choice that weighs the costs and benefits of being formal or informal (Bruhn and McKenzie 2014; Perry et al. 2007). The growing more recent evidence on the effects of formalization policies supports such a view (Benhassine et al. 2018; de Mel et al. 2013). These studies have yielded important insights into the typically limited efficacy of formalization policies to increase formalization rates and improve firm performance.

Our study adds a complementary perspective to the literature on the characteristics and performance of informal firms, selection into formality, and formalization policies. We look at the case of Viet Nam, a developing economy where firms have been formalizing faster than elsewhere. To our knowledge, this paper is the first to examine productivity dynamics of, misallocation among, and aggregate productivity contributions of formal, informal, and formalizing firms. We rely on long-term observational data; thus, we examine endogenous selection or transition into formality, although (exogenous) policy changes affect formalization. We apply extensions of the decomposition methods of Olley and Pakes (1996) (henceforth, OP) and Melitz and Polanec (2015) (henceforth, MP) to a unique firm-level panel dataset on private small- and medium-sized enterprises (SMEs) in the manufacturing sector in Viet Nam that covers a period of ten years. In this period, we observe an increasing and recently accelerated formalization of firms driven by the formalization of small household businesses. This process and acceleration cannot be clearly attributed to one specific policy change, but a recent law that penalizes non-compliance with registration obligations for household businesses is likely to (partly) explain the recent acceleration of formalization among those firms. With the data at hand, we cannot and do not directly study the impact of formalization policies. Building on and extending the OP and MP frameworks, we examine the role of informal–formal transitions together with firm entry and exit for formal and informal sector productivity. Our decomposition analysis quantifies the contributions of these firm dynamics to productivity growth and assesses the productivity dispersion among the different groups of firms—in particular formalizers, formal survivors, and informal survivors.

We find that formalizing firms in Viet Nam make an important contribution to aggregate productivity growth among manufacturing SMEs. Formalizing firms grow much faster than other firms and formalization is associated with increasing efficiency: among formalizing firms, firms with higher total factor productivity (TFP) gain market shares. Yet, growth of formal survivors that account for the largest share of production is the main driver of overall productivity growth in SMEs. We can identify two different ‘regimes’ of formalization. Until early 2010 more

productive (previously) informal firms formalize, indicating endogenous selection of better-performing informal firms. Policy changes and accelerated formalization then alter the characteristics of formalizers, as less productive firms also become formal. In this period, labour productivity of formalizers stagnates while TFP increases sharply, suggesting important adjustments among these firms. While this formalization wave depresses average formal TFP growth, the overall productivity effect is positive.

The remainder of the paper is organized as follows. In Section 2, we provide a brief overview of how our work relates to the existing literature. In Section 3, we describe the data, and present descriptive statistics as well as the production function estimates. Section 4 focuses on formalization and firm dynamics. In Section 5, we explain and apply the productivity decomposition and report the results. Section 6 concludes.

2 Related literature

This paper relates to and combines insights from various strands of literature. As we are interested in the formal and informal firm dynamics that drive aggregate productivity, we build analytically and methodologically on the productivity decomposition approach originally proposed by Olley and Pakes (1996). This approach decomposes aggregate productivity into growth that comes from average productivity gains or from reallocation gains that result from a higher share of output being produced by more productive firms (or plants, in Olley and Pakes 1996). The decomposition was further developed by Melitz and Polanec (2015) to explicitly account for the contributions of surviving, entering, and exiting firms. The OP and MP approaches have typically been applied to study changes in aggregate productivity and the evolution of productivity dispersion over time within the same country (and the same sector). Our work is similar to that of Du et al. (2014) who compare the contributions of state and non-state firms to aggregate TFP growth in Chinese manufacturing from 1998 to 2007. The study applies the MP decomposition to state and non-state firms. It finds that the observed slowdown in aggregate productivity growth is driven partly by a worsening resource misallocation within the state sector.

Another approach to examining resource misallocation follows the work of Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) (henceforth, HK). These studies establish ‘efficient counterfactuals’ to quantify the aggregate productivity losses from misallocating resources across farms and firms.¹ Hsieh and Klenow (2009) investigate the extent of factor misallocation in China and India. Their study suggests that the countries could boost their TFP by 40–60 per cent and 30–50 per cent, respectively, if the extent of dispersion of marginal products were identical to the level in the United States. In a study of land allocation across farms in Malawi, Restuccia and Santaaulalia-Llopis (2017) find that aggregate agricultural output could be up to four times higher if factors were allocated to their efficient use among existing farmers.

Recent studies have warned against equating dispersion to inefficiency. Productivity dispersion among firms can also be driven by adjustment costs and uncertainty, market power, quality, and technology; that is, factors that may also positively affect aggregate productivity (Cusolito and

¹ Hsieh and Klenow (2009) rely on a model of monopolistic competition, in which producers face competitive labour and capital markets. Under these assumptions, differences in marginal products of labour and capital across firms or plants (ideally, within narrowly defined industries) reflect factor misallocation.

Maloney 2019).² For agriculture, Gollin and Udry (2019) highlight measurement error, heterogeneity in unobserved land quality, as well as (late-season) production shocks as sources of ‘efficient’ productivity dispersion. Using the productivity dispersion between plots within farms as an efficient benchmark, they show that the traditional HK approach applied, for example, by Restuccia and Santaaulalia-Llopis (2017) may overestimate misallocation by a factor of 2 to 4. This exercise still leaves ‘non-trivial’ (Gollin and Udry 2019) inefficient dispersion that is well below earlier estimates.

With these caveats in mind, studies based on the HK framework still offer valuable insights although they may overestimate the extent of misallocation of production factors across firms. A few of those studies explicitly examine the informal sector, including Busso et al. (2012) and Levy and López-Calva (2018) who study the Mexican case. Both studies document productivity, size, and misallocation distributions for formal and informal firms and find that informal firms are much less productive than formal firms; yet, they command a large share of resources and hence contribute considerably to low productivity in Mexico. Levy and López-Calva (2018) find that in the absence of misallocation average earnings would be higher and earnings differentials across schooling levels would widen, raising the returns to education.

Indicative of resource misallocation related to informality are also the many studies that have shown relatively high marginal returns to capital in microenterprises (Blattman et al. 2016; Grimm et al. 2011; McKenzie and Woodruff 2008). However, credit market imperfections partly explain this finding that does not universally hold for all microenterprises. In particular, some studies suggest that returns may be lower in subsistence-type activities (Fafchamps et al. 2014; Grimm et al. 2012). While (many) informal firms may be more affected by capital market distortions than formal firms, quite a number of potentially distortionary policies may only apply to formal firms. In a recent paper on Viet Nam, McCaig and Pavcnik (2018) show that the removal of an output market distortion that is more binding for initially more productive firms, such as a tariff on exports, induces a movement of workers away from less productive employers in informal microenterprises to employers in the more productive formal enterprise sector.

The heterogeneity of informal firms is likely to reinforce the ambiguities in the relationship between productivity and formalization. Potential formalization benefits for firms could arise from improved access to credit, greater scope for marketing and engaging with larger firms and the government, less risk of fines, and better access to (government) training and support programmes (Bruhn and McKenzie 2014). Yet, formalization also comes along with being subject to taxation and regulation and, potentially, increased exposure to corrupt officials. This ambiguity can explain the available evidence that suggests that formalization benefits are, on average, non-existent or negligible at best (Benhassine et al. 2018; Bruhn and McKenzie 2014; de Mel et al. 2013). These average effects may hide important heterogeneity that is only identified in a few studies (e.g., McKenzie and Sakho 2010). Such heterogeneity of formalization should be expected given the heterogeneity of informal firms as well as important differences in (investigated) formalization policies, ranging from increased enforcement to formalization subsidies.

² A dynamic extension of the HK framework by David and Venkateswaran (2019) tries to address these concerns by explicitly accounting for adjustment costs and uncertainty in the firm’s investment decision. For a number of developing countries, David et al. (2019) show that technology heterogeneity explains an important part of productivity dispersion, but it turns out that adjustment costs and informational frictions only play a minor role. This clearly indicates that dispersion cannot be equated with inefficiency, but in many cases the exercise still leaves half of the dispersion or more unexplained.

Ambiguous evidence on the effects of formalization has also been presented for Viet Nam. Applying a matched double-difference approach to two waves of the survey (2007 and 2009) used in this paper, Rand and Torm (2012) find formalization to increase profits for SMEs by about 15 per cent and to decrease the share of casual workers employed by the firm. Using household business and informal sector surveys (with businesses identified from a labour force survey) for 2007 and 2009, a similar quasi-experimental study by Demenet et al. (2016) comes to similar conclusions. They find formalization to be associated with a 20 per cent increase in value added, except for very small firms. These gains appear to be driven by improved access to better locations (indoor premises) and new equipment (electricity and internet) and better business practices (keeping written accounts). Note that manufacturing firms only account for about 20 per cent of the sample in this dataset. In contrast, a similar study by McCaig and Nanowski (2019) finds that formalization does not lead to any improvements in business performance. They rely on data on non-farm household businesses from three waves (over four years, 2004–08) from the Viet Nam Household Living Standards Survey (VHLSS) and argue that their results differ from studies because they can control for pre-formalization trends. The mixed evidence on formalization effects is in line with qualitative and quantitative evidence presented by Cling et al. (2012) on the perceptions of informal household businesses. These business owners often perceive registration procedures as complicated and taxes as arbitrary. They see some formalization benefits, such as less exposure to corruption, but fear the administrative burden.

These studies look at the effects of formalization before the more recent acceleration of formalization rates after 2013. They correctly note that Viet Nam even in the late 2000s is a relatively ‘formalized economy’ given its income level (e.g., see McCaig and Nanowski 2019) that, in the period considered by the above-mentioned empirical studies, sees a further steady increase in formality. The informal firm dynamics in this period are described by McCaig and Pavcnik (2017), again using the three waves of the VHLSS. They find business exit to be prevalent in the informal sector driven by less productive firms. Yet, these firms are replaced by equally unproductive entrants that typically have low growth prospects. Their analysis suggests that only very few informal firms transition into large, more successful (and productive) firms.

3 Data and production function estimates

We use the Viet Nam SME database, a survey of small and medium-sized non-state enterprises operating in the manufacturing sector, compiled in a collaborative effort by the Central Institute for Economic Management, the Institute of Labour Science and Social Affairs, the Development Economics Research Group at the University of Copenhagen, and United Nations University World Institute for Development Economics Research.³ Between 2005 and 2015, the data were collected biennially and consist of three modules. First, a main enterprise module for owners or managers. Second, an economic accounts module collected alongside the main enterprise module and containing annual data on key firm-level performance/economic indicators (on the past two years; that is, these indicators are available annually between 2004 and 2014). Third, an employee module of a random subset of employees in a quarter of randomly selected enterprises, collected biennially since 2007.

About 2,500–2,800 enterprises are sampled from ten selected provinces, which jointly covered 30 per cent of the country’s overall non-state manufacturing population at the beginning of the survey

³ Detailed information about the surveys, the sampling methodology, and the data can be obtained from CIEM et al. (2005–15) and DERG (2005–10).

in 2005, and surviving firms were re-interviewed. Over the whole sample period between 2005 and 2015, a total of 5,134 firms were surveyed, resulting in 15,810 firm-year observations. Firms are of nine different types of ownership, the largest groups being household businesses (65 per cent), limited liability companies (19.95 per cent), and sole proprietorship firms (8.08 per cent).⁴ Exiting firms were replaced with randomly selected new firms holding constant the level of household businesses from the 2002 Establishment Census (GSO 2004) and of the updated population of firms registered under the Enterprise Law in the years in which the surveys took place (Rand and Tarp 2020).

To limit firm heterogeneity and to make more comparable formal and informal firms, we restrict our analyses to firms with 50 or fewer employees, thus dropping 518 observations.⁵ Furthermore, we focus on private firms, ending up with 14,511 observations in total. Of the overall 4,485 firms in our reduced sample, household businesses account for 66 per cent, limited liability companies for 19 per cent, sole proprietorship firms for 9 per cent, joint stock companies for 3 per cent, and collectives for 3 per cent—so the distribution of firm categories is similar to the full sample. Note that household businesses make up 97 per cent of the firms in Viet Nam and that their number almost doubled over our sample period 2005–15, from 2.7 to 4.7 million (Rand and Tarp 2020).

The registered enterprise sector consists of limited liability companies, sole proprietorships, and collective and foreign enterprises. Being a registered enterprise under the Enterprise Law requires a firm to adhere to accounting procedures and report on the firm's financial position and on the employed workforce. Further, these firms are all registered with the tax authorities. What separates the registered enterprise sector from the household business sector is a size threshold of ten regular employees. Household businesses with fewer than ten employees are required by law to register their business with some exceptions granted to small-scale services with low income (defined by local authorities).⁶ They are also required to register with the tax authorities at the local tax office within 10 days of registering the business (Rand and Tarp 2020). Thus, formal household businesses should have (i) a business registration licence, and (ii) a tax code number. While there are a few registered household businesses without a tax code, essentially all household businesses with a tax code are also registered.⁷ We define a household business as informal if it does not possess a tax code number.⁸

⁴ The remaining ownership categories were partnership (0.27 per cent), collective/cooperative (2.76 per cent), joint stock company with state capital (0.25 per cent), joint stock company without state capital (3.34 per cent), joint venture with foreign capital (0.01 per cent), and state enterprise (local) (0.04 per cent).

⁵ However, we keep all observations of firms who over the course of the survey period were below and above the size threshold to prevent these firms from being mistaken as exiters.

⁶ See Government of Viet Nam (2006). Because the threshold refers to regular employees some household businesses may have more than ten workers by the time they are surveyed.

⁷ While there were different registration procedures for the business registration license and the tax code until 2010, in June of that year the government introduced the so-called enterprise code number that combines the business license and the tax code.

⁸ We are, of course, aware that there are various other ways how informality can be defined, and that some recent studies stress that informality is better understood as a continuum with varying degrees (Williams et al. 2016; Ulyssea 2018).

Among household businesses we observe a substantial formalization between 2005 and 2015 that drives the overall formalization of manufacturing firms. Table 1 shows that the share of firms with a tax code number increases from 55.1 per cent in 2005 to 89.5 per cent in 2015. We observe an accelerated formalization of firms in more recent years; that is, by about 20 percentage points between 2013 and 2015.

Table 1: Number of firms with a tax code number over time

Tax code	Year												
	2005		2007		2009		2011		2013		2015		Total
	No.	Col %	No.	Col %	No.	Col %	No.	Col %	No.	Col %	No.	Col %	Col %
No	1,035	44.9	1,003	40.8	896	35.8	702	29.7	707	29.6	263	10.5	31.7
Yes	1,269	55.1	1,456	59.2	1,606	64.2	1,659	70.3	1,685	70.4	2,230	89.5	68.3
Total	2,304	100.0	2,459	100.0	2,502	100.0	2,361	100.0	2,392	100.0	2,493	100.0	100.0

Source: authors' calculation based on UNU-WIDER's Viet Nam small- and medium-sized enterprise (SME) database (see CIEM et al. 2005–15).

The accelerated formalization of household businesses is likely to be a reaction to the 2013 decree on penalties on administrative violations in commercial activities, including violations with regard to business registration certificates and business licences. Specifically, the law stipulates fines between 5,000,000 and 10,000,000 Vietnamese dong (approximately US\$200–400) for 'acts of business operation without business registration certificate in accordance with regulations' (Government of Viet Nam 2013).

The rising share of formal household businesses leads to a decline in informal employment (see upper panel of Figure 1). While in 2005 about a third of manufacturing employment in Viet Nam was informal, it had almost disappeared by 2015. This trend is more continuous than the share of registered enterprises as even before the new regulations of 2013 informal household businesses employed fewer workers per firm. At the same time, the average number of workers in formal firms fell, as the smaller household businesses formalized, in particular after 2013. The lower panel of Figure 1 shows the associated trends in output and in labour productivity (output per worker) in formal and informal firms. We observe that overall output rises due to increasing labour productivity with the already low contribution of informal household businesses falling from a small share (less than 10 per cent) in 2005 to an almost negligible share in 2015. The figure also illustrates the huge productivity differences between formal and informal firms: with an initial log difference of 1.5, formal firms are about four to five times more productive than informal firms. This gap narrows over time, as average labour productivity among informal firms grows faster than among formal firms. And while the sharp increase in formality in the most recent year is even associated with a slight drop in average formal sector labour productivity, it further increases among the household businesses that decide to remain informal.

Figure 1: Trends in employment and output and labour productivity



Note: total output is measured as value added in million Vietnamese dong and total employment as the number of full-time works in thousands.

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Appendix Table A1 depicts how formal and informal firms are distributed across two-digit manufacturing International Standard Industrial Classification (ISIC) sectors using the pooled (and reduced) sample. There are quite a few industries in which formal and informal firms coexist, including food/beverages, textiles/leather/apparels, wood/furniture, metal products, and non-metallic mineral products (bricks, tiles, etc.), with informal firms clearly outnumbering formal ones in food/beverages and wood/furniture. There are couple of sectors that are mostly populated by formal firms such as paper, publishing/printing, electronic machinery, and the manufacturing of rubber. With the exception of food/beverages, which becomes more important in the informal sector, the industry composition changes little over time among formal and informal firms. An exception is the last period. Here, the entry of previously informal firms causes an increase in the number of formal firms in food/beverages as well as in wood manufacturing.

Although formal and informal firms often operate in the same sectors, there are marked differences between them with respect to both firm-level and owner characteristics, which we report for both types of firms in Table 2.

Table 2: Firm and owner characteristics

	Firm characteristics						Owner characteristics			
Formal	Sales	Capital	No. employees	Formal accounting	Credit constrained	Age	General secondary	Professional secondary	Age	Observations
No	11.9	11.5	4.1	0.0	0.20	13.1	0.41	0.04	44.5	1,758
Yes	13.6	13.3	14.4	0.6	0.24	10.5	0.72	0.31	43.3	2,727

Note: characteristics are used from the first time when a firm was surveyed. Sales and capital stocks are measured in logs.

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

In line with the previous observations, formal firms record noticeably higher sales, capital, and numbers of employees than informal firms. Sales of formal firms are five to six times higher, and formal firms employ six times more capital. They employ 14 regular workers compared with 4 on average in informal firms. Formal firms are on average 2.5 years younger than their informal counterparts, and their owners are slightly younger. As expected, there are virtually no informal firms that maintain a formal accounting book in accordance with government guidelines, but the share is also relatively low for formal firms with only 61 per cent. This rather low share is attributable to the fact that only about 16 per cent of formal household businesses maintain formal accounting. When looking at (perceived) access to credit the data suggest that a somewhat larger share of formal firms (about a quarter) is credit constrained.⁹ Informal firms may have lower demand for credit to start with and thus are perceived being less constrained.¹⁰ Very striking differences can be seen in the education of the owner: while 72 per cent (30 per cent) of formal firm owners have at least finished upper secondary education (professional secondary education), this is the case for only 41 per cent (0.04 per cent) of informal firm owners.

Such differences in human and possibly managerial capital are likely to be associated with higher TFP in formal firms. We back out TFP by estimating the input coefficients in a value added Cobb–Douglas production function. In Appendix Table A2 we report labour and capital elasticities for the pooled sample as well as average TFP that we also compute for informal and formal firms separately—using ordinary least squares (OLS) and the approach proposed by Akerberg et al. (2007) (henceforth, ACF). Further, we pool observations across the two-digit ISIC sectors depicted in Appendix Table A1 as for a sector-level and a separate estimation for formal and informal firms we do not have enough observations. In Appendix B, we also provide details on the production function estimates and the different estimation procedures. Note that the labour and capital elasticities do not differ significantly between formal and informal firms—also not when estimating them for single sectors where the two types of firms have larger overlap.

Indeed, average TFP is higher in formal than in informal activities, in a range between 30 and 80 per cent depending on the estimation model. Our preferred ACF estimation puts the average TFP advantage of formal over informal firms at roughly 80 per cent. With around 0.15 the capital coefficient lies in a range that is commonly found in production function estimates using firm-level data. The labour coefficient for the pooled sample is slightly below 1 using the ACF estimation (slightly above 1 using OLS and around 0.7 using fixed effects). This suggests slightly increasing returns to scale in our sample, in line with previous production function estimates using

⁹ Following the definition in Kinghan et al. (2018), firms are defined as credit constrained if they are either credit rationed (i.e. they were rejected for a formal loan) or if they did not apply for a loan because of a possible rejection, too difficult of a process, or too high interest rates.

¹⁰ This is in line with Kinghan et al. (2018) who find for the same sample of Vietnamese small and medium-sized enterprises that rejected credit applications are limiting investment activity, particularly for firms with higher investment efficiency which are arguably more likely to be formal.

the same data (Rand and Tarp 2020). Notice that the TFP measure we generate following the ACF procedure is based on a production function in revenues and not in physical quantities, and is thus better described as firm profitability.¹¹ We have data on physical output quantities of the top three products of a firm. Unfortunately, product units are measured imprecisely, without product-specific codes, introducing too much noise in a quantity-based production function estimation.¹²

4 Formalization and firm dynamics

We now explore in detail firm dynamics in the manufacturing sectors in Viet Nam, including not only informal–formal transitions but also entry and exit. Table 3 depicts the number and share of firms transitioning between, entering into, or exiting from formal and informal sectors over the entire survey period, 2005–15. Note that the sampling procedure leads to an underrepresentation of entrants, as replacements for exiting firms are drawn randomly from the entire registered enterprise and household business (manufacturing firms) populations while leaving all surviving firms in the sample, which will thus grow older (compared with the population) over time. We define entrants according to the time horizon covered by the respective analysis. For example, in an analysis that covers 2005–15 all firms that enter the market in these ten years are considered entrants, but not those born before 2005. Similarly, firms are classified as formal survivors when they report to have been registered at birth and informal survivors if not.¹³ The sample does not contain any firms that are lost due to attrition so that firms that exit the sample can be considered ‘true’ exiting firms.

The first panel of Table 3 reports transitions between the formal and informal sector as well as exits and entries for firms that were sampled in 2005, in 2015, or in both periods. The second panel further breaks down movements for the subset of firms that did not appear in the 2005 sample, but existed in 2015. We split these firms into three groups: (i) informal at start-up; that is, firms that had been informal when started and in operation before 2005; (ii) formal at start-up; that is, firms that had been formal when started and in operation before 2005; and (iii) new entrants; that is, firms that started business operations since 2005.

The panel-data-based transitions confirm that a large share of informal firms formalized. Of the 1,035 informal firms that were part of the sample in 2005, 441 or 42.6 per cent had formalized until 2015 (first row of the first panel in Table 3). Similarly, of the 460 initially informal firms (informal at start-up) that had not been part of the 2005 survey, 74.6 per cent had formalized. Further, new entrants are an important driver of formalization. By 2015, 84.2 per cent of the entrants in the past ten years are formal.

¹¹ Syverson (2004) and Foster et al. (2008) were among the first to show that the difference between physical total factor productivity (TFPQ) and revenue total factor productivity (TFPR) may, in some applications, be critical as the traditional TFPR measure of revenue over costs may conflate firm efficiency with other demand-side factors. For an excellent discussion, see Cusolito and Maloney (2019).

¹² Using data on outputs in physical units, De Loecker et al. (2016) show how an extension of the ACF framework may yield consistent estimates of the production function coefficients, offering treatment of two additional biases that are not addressed within the standard ACF framework, stemming from the unobserved allocation of inputs across products within multi-product firms, i.e. the input allocation bias, and the bias stemming from unobserved input prices (due to the use of quality-differentiated inputs) by firms (i.e. the input price bias).

¹³ This information is available only for the last two survey waves, 2013 and 2015. When it is not available, we assume the current registration status is also the one at birth.

Table 3: Transitions, 2005–15

	Status 2015											
	Informal			Formal			Exit			Total		
	No.	Row %	Col. %	No.	Row %	Col. %	No.	Row %	Col. %	No.	Row %	Col. %
Status 2005												
Informal	26	2.5	9.9	441	42.6	19.8	568	54.9	43.2	1,035	100.0	27.2
Formal	8	0.6	3.0	515	40.6	23.1	746	58.8	56.8	1,269	100.0	33.3
Not in sample	229	15.2	87.1	1,274	84.8	57.1	0	0.0	0.0	1,503	100.0	39.5
Total	263	6.9	100.0	2,230	58.6	100.0	1,314	34.5	100.0	3,807	100.0	100.0
Not in sample												
Informal at start-up	117	25.4	51.1	343	74.6	26.9	0	0	0	460	100.0	30.6
Formal at start-up	10	2.5	4.4	388	97.5	30.5	0	0	0	398	100.0	26.5
New entrant	102	15.8	44.5	543	84.2	42.6	0	0	0	645	100.0	42.9
Total	229	15.2	100.0	1,274	84.8	100.0	0	0	0	1,503	100.0	100.0

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Over the whole survey period, most formal and informal firms exited and ceased operations altogether. At 58.8 per cent (over ten years) compared with 54.9 per cent, the exit rate is somewhat higher for formal than for informal firms. High formalization and exit rates together imply that only very few informal firms survived when they remained informal. We can also see that the surviving initially formal firms basically all remained formal. Further examination of the trajectories of formal firms that become informal (not reported) suggests that these firms tend to eventually exit.

The long time horizon conceals important dynamics including the different speeds of formalization. This is why we examine formal–informal transitions, entry into, and exits from formal and informal sectors considering shorter timespans of two years, corresponding to the periods between two survey waves. We now define new entrants as those firms that started business operations within the two years before the interview took place. Note again that we know whether firms were registered at business start-up. An added advantage of adopting shorter time horizons is that we now also include firms that were neither surveyed in 2005 nor surveyed in 2015.

Splitting the ten-year period into the five biannual periods (2005–07, 2007–09, 2009–11, 2011–13, and 2013–15) corroborates many of the previous findings, while offering some new insights (see Appendix Table A3). First, the formalization rate is considerable. About 15–18 per cent of informal firms formalize within two years in the first three periods. Between 2011 and 2013 the formalization drops to 7.7 per cent and then jumps to 80.3 per cent between 2013 and 2015. Further, with 20.4–37.2 per cent formalization rates for firms that were not part of the previous survey and had been informal at business start-up (first rows of the second panels of Appendix Table A3 are always somewhat higher than those for informal incumbents. Second, across all periods, exit is around 15–20 per cent (with the exception of 2005–07) and larger in the formal than in the informal sector, by around 3–5 percentage points. Third, as indicated in the third rows of the second panels of Appendix Table A3, most new entrants are registered when beginning business operations. In more recent years, however, the share of informal entrants almost doubled. While in the periods 2005–07, 2007–09, and 2009–11 the share was between 16.2 and 20 per cent, it increased to 34.3 per cent in 2013 and to as much as 40.5 per cent in 2015. So while existing firms quickly formalize, firms continue to enter without registration. In sum, both the short and the long time horizons suggest that informal–formal transitions are common, formal–informal

transitions are rare, and that firm entries and exits are important. Hence, to examine changes in productivity in both sectors and at the aggregate level, accounting for these firm dynamics is critical.

Before laying out our framework that allows us to explore these different margins and their contributions to productivity growth more carefully, we depict in Table 4 firm and owner characteristics for the distinct types of firms depicted in the transition tables (i.e. Table 3 and Appendix Table A3): survivors in the informal and formal sector, exiters from the informal and formal sector, new entrants to the informal and informal sector, and firms transitioning between the two sectors (i.e. informalizers and formalizers).¹⁴

Table 4: Characteristics of survivors, exiters, entrants, and (in)formalizers

	Firm characteristics						Owner characteristics			
	Sales	Capital	No. employees	Formal accounting	Credit constrained	Age	General secondary	Professional secondary	Age	Observations
Survivor										
In informal	11.7	11.4	3.5	0.00	0.17	17.8	0.36	0.02	46.4	1,489
In formal	13.6	13.4	15.2	0.56	0.21	14.6	0.74	0.16	46.4	5,904
Transitioning firm										
Informalizer	12.5	12.5	5.5	0.04	0.23	16.9	0.52	0.06	47.9	216
Formalizer	12.2	11.8	4.7	0.07	0.18	17.5	0.53	0.04	47.6	1,106
Complete exiter										
From informal	11.7	11.5	3.7	0.02	0.16	16.7	0.42	0.03	48.0	605
From formal	13.5	13.3	12.2	0.60	0.23	12.0	0.73	0.21	45.1	1,387
New entrant										
To informal	11.7	10.9	3.6	0.01	0.26	2.5	0.48	0.06	39.1	199
To formal	13.5	12.7	14.5	0.71	0.28	2.5	0.79	0.28	38.6	506

Note: sales and capital stocks are measured in logs.

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Mirroring the findings of Table 2, formal survivors record considerably higher sales, capital, and numbers of employees. Further, their owners are much more educated in terms of both general and professional educational attainment than formalizers. Compared with informal firms, formalizers (and informalizers) record higher sales, capital, and numbers of employees and their owners are much more educated, but the opposite is true when comparing them with formal firms. So these transitioning firms really fall in between incumbent formal and informal firms. Both formal and informal exiters resemble incumbents of their respective sector more closely than one may anticipate. Formal survivors are somewhat larger and older on average than exiters. Formal and informal entrants are also fairly similar to incumbents in their respective sector. As expected, both formal and informal entrants have smaller capital stocks and formal entrants tend to have a smaller number of employees. Yet, formal entrants are more likely to maintain formal accounts and their owners are more educated than informal incumbents.

Finally, we break down the group of formalizing firms into firms formalizing before 2013 and those formalizing between 2013 and 2015 when the formalization rate roughly quadrupled to 80.4 per cent. In our sample, about half of the formalizers over the entire period 2005–15 formalize

¹⁴ Here, we consider only those firms that participated in the survey when the transition occurred—hence disregarding firms for whom much time has elapsed between the transition and the time that they were surveyed—since we wish to examine the characteristics of the firms right when they transitioned.

between 2013 and 2015. We expect the acceleration of formalization to have had an impact on the characteristics of formalizers, which is confirmed by Table 5. At the time of formalizing, early formalizers (before 2013), have much higher sales, are considerably better capitalized, and have four employees more than formalizers of 2013–15. The later formalizers also keep formal accounts much less frequently and are considerably older, as are their owners.

Table 5: Characteristics of formalizers before and after 2013

Transition period	Firm characteristics						Owner characteristics			
	Sales	Capital	No. employees	Formal accounting	Credit constrained	Age	General secondary	Professional secondary	Age	Observations
2005–13	12.8	12.7	6.7	0.14	0.22	14.6	0.53	0.04	46.4	531
2013–15	11.7	11.0	2.8	0.01	0.14	20.2	0.52	0.04	48.8	575

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

5 Decomposition analysis

Our framework to examine the contribution of informal–formal transitions and of entry and exit builds on the OP and MP productivity decompositions, but contains additional terms capturing the contribution of transitions between the formal and informal sector to productivity growth.

5.1 The OP decomposition and application

Following Olley and Pakes (1996), we measure the (log) aggregate (average) productivity level Φ_t in each period as the weighted average of (log) plant-level productivity, with shares of aggregate output as weights

$$\Phi_t = \sum_i^N s_{it} \varphi_{it},$$

where Φ_t is aggregate productivity at time t , φ_{it} is the productivity of firm i at time t , and s_{it} is i 's share in aggregate output at time t . The main insight of the OP decomposition was then to show that the weighted productivity average can be decomposed into the following two separate terms:

$$\Phi_t = \bar{\varphi}_t + \sum_i (s_{it} - \bar{s}_t)(\varphi_{it} - \bar{\varphi}_t)$$

$$\Phi_t = \bar{\varphi}_t + \text{cov}(s_{it}, \varphi_{it}),$$

where $\bar{\varphi}_t = \sum_{i=1}^{n_t} \varphi_{it} / n_t$ is the unweighted average productivity and $\bar{s}_t = 1/n_t$ is the average market share. Hence, the second term in the first equation captures the (summed) joint variability of firms' productivity deviation (from mean productivity) and firms' market share deviation (from mean market share), which is essentially the covariance between a firm's productivity and firm's market share. Consequently, changes in productivity over time, Φ_t , can be attributed to changes in the unweighted mean $\bar{\varphi}_t$, which constitute shifts in the productivity distribution, and/or changes in the covariance $\text{cov}(s_{it}, \varphi_{it})$, which captures market share reallocations. Table 6 depicts

the result of the OP decomposition for (log) TFP and (log) labour productivity, respectively.¹⁵ We apply the decomposition for all firms as well as for the formal and the informal sector separately, and trim the 0.5 per cent tails of TFP and the market share in both sectors to make the results robust to outliers.¹⁶

The strong growth in overall labour productivity is to an important extent driven by strong growth in TFP. Cumulatively, weighted TFP grows by 97 per cent and labour productivity by 99 per cent over the ten years under consideration. Among all manufacturing SMEs, the TFP covariance terms vary over time with no clear trend. For labour productivity, the overall covariance term increases indicating that firms with higher labour productivity gain market share.

Table 6: OP decompositions

Year	All				Formal				Informal			
	Weigh.	Unweigh.	Cov.	Obs.	Weigh.	Unweigh.	Cov.	Obs.	Weigh.	Unweigh.	Cov.	Obs.
Log TFP												
2005	8.49	8.01	0.48	2304	8.61	8.39	0.22	1269	7.61	7.54	0.07	1035
2007	8.80	8.25	0.55	2459	8.89	8.61	0.27	1456	7.96	7.72	0.23	1003
2009	8.84	8.34	0.50	2502	8.91	8.68	0.23	1606	7.90	7.75	0.15	896
2011	8.95	8.51	0.44	2361	9.01	8.78	0.23	1659	7.99	7.87	0.12	702
2013	9.06	8.58	0.49	2392	9.12	8.82	0.30	1685	8.10	7.99	0.11	707
2015	9.17	8.72	0.46	2493	9.19	8.78	0.41	2230	8.58	8.22	0.36	263
Log labour productivity												
2005	9.76	9.61	0.15	2304	9.87	9.87	-0.00	1269	9.33	9.29	0.04	1035
2007	10.09	9.89	0.21	2459	10.20	10.16	0.04	1456	9.54	9.49	0.05	1003
2009	10.20	9.93	0.27	2502	10.28	10.18	0.09	1606	9.64	9.49	0.15	896
2011	10.40	10.12	0.28	2361	10.45	10.31	0.14	1659	9.92	9.67	0.25	702
2013	10.38	10.12	0.26	2392	10.43	10.31	0.12	1685	9.84	9.67	0.17	707
2015	10.43	10.15	0.27	2493	10.44	10.20	0.25	2230	9.89	9.77	0.12	263

Note: Weigh., weighted; Unweigh., unweighted; Cov., covariance; Obs., observations; TFP, total factor productivity.

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Weighted TFP is much higher in the formal sector with a gap of 1 log point until 2013. In the last period, the gap shrinks to 0.6 due to strong growth in informal TFP. Between 2013 and 2015, unweighted formal sector TFP even declines by 4 log points. The trends are similar for labour productivity. Labour productivity is higher in formal than in informal firms but the gap between the two sectors fluctuates. It widens in earlier years, but narrows between 2013 and 2015. In this period, unweighted formal labour productivity drops, while informal labour productivity (both weighted and unweighted) rises. This drop is likely to be driven by formalizing firms with substantially lower labour productivity than formal sector incumbents. The increase of informal labour productivity, in turn, is due to the strong increase in informal TFP. During the formalization wave (i.e. between 2013 and 2015), bigger informal firms with higher capital endowments—relative

¹⁵ Following Melitz and Polanec (2015), in the case of labour productivity as the productivity measure, we use employment shares as weights.

¹⁶ Note that besides the trimming exercise, we lose additional observations for which TFP could not be computed because either data on value added, capital stocks, or the number of employees was missing, ending up with 14,012 observations in total.

to remaining informal sector firms—may move to the formal sector and not necessarily informal firms with higher TFP.

The covariance term for formal TFP is relatively constant in earlier years, but is highest in the last period; that is, the increase in formality is associated with a rising market share of relatively more productive firms within the formal sector. The increase in the covariance term between 2013 and 2015 is very pronounced for formal labour productivity. This is plausible given that relatively small previously informal firms enter formality. These firms are less productive than formal incumbents and have low market shares (among formal firms). Misallocation—measured by the covariance term for TFP—is consistently higher (the covariance lower) for informal than for formal firms. In the last period, the covariance term for TFP increases while it falls for labour productivity. Accelerated formalization, thus, does not appear to leave behind only the least productive firms, as firms with higher TFP account for higher shares of value added. The simultaneous drop in the labour productivity covariance terms suggests that the informal survivors are relatively more labour-intensive.

In sum, the decomposition results are indicative of a selection process into formality that has been altered by the policies that were put in place. They have changed not only the speed, but also the characteristics of firms that have moved into formality. Accordingly, they clearly illustrate the key role of informal–formal dynamics for understanding the development of aggregate productivity and the dispersion of productivity between firms. From this static decomposition alone, however, we are unable to tell how much of the patterns that we observe in weighted and average productivity and the covariance between productivity and market shares are driven by informal–formal transitions and how much may in fact be attributable to new firms entering into and existing firms exiting from the formal and informal sectors. Therefore, we now explicitly assess the role of these dynamics as well as of entry and exit.

5.2 The extended MP decomposition and application

We extend the decomposition proposed by Melitz and Polanec (2015) and accommodate (i) entries to the formal/informal sector from the informal/formal sector, (ii) and exits from the formal/informal sector to the informal/formal sector, adding two additional terms to the decomposition. In the first of two periods, we add an additional term for exiters and in the second period for entrants. Considering the formal sector and two periods, its aggregate productivity level in the first period is thus a function of the aggregate share and aggregate productivity of formal survivors, exiters quitting business operations, and exiters to the informal sector. In the second period, we consider formal survivors, new entrants that start a formal business, and formalizers, (i.e. entrants from the informal sector). Accordingly, the five groups contributing to aggregate productivity of the formal sector are formal survivors (SF), exiters to the informal sector (informalizers, XI), complete exiters because of firm death (complete exiters, XT), entrants from the informal sector (formalizers, EI), and entrants who start business operations (new entrants, EN). Supposing two periods, only the group of formal survivors is observed in both periods, whereas the exiters and the entrants are only observed in the first or second period, respectively. The following equations show how these five groups contribute to formal sector aggregate productivity in the first and second periods:¹⁷

$$\Phi_1 = s_{SF1}\Phi_{SF1} + s_{XI1}\Phi_{XI1} + s_{XT1}\Phi_{XT1}$$

¹⁷ The same logic applies to the informal sector. We will, however, refrain from a detailed decomposition of the informal sector dynamics because transitions to the informal sector from the formal sector are much less common.

$$\Phi_1 = \Phi_{SF1} + s_{XI1}(\Phi_{XI1} - \Phi_{SF1}) + s_{XT1}(\Phi_{XT1} - \Phi_{SF1})$$

$$\Phi_2 = s_{SF2}\Phi_{SF2} + s_{EI2}\Phi_{EI2} + s_{EN2}\Phi_{EN2}$$

$$\Phi_2 = \Phi_{SF2} + s_{EI2}(\Phi_{EI2} - \Phi_{SF2}) + s_{EN2}(\Phi_{EN2} - \Phi_{SF2}),$$

where $s_{Gt} = \sum_{i \in G} s_{it}$ represents the aggregate market share of a group G of firms ($G = \{SF, XI, XT, EI, EN\}$), and $\Phi_{Gt} = \sum_{i \in G} (s_{it}/s_{Gt})\phi_{it}$ is that group's (weighted) aggregate productivity. Φ_1 , the (weighted) aggregate productivity in the first period, is the sum of the (weighted) aggregate productivity of the formal survivors, the informalizers, and the complete exiters. Similarly, Φ_2 , the (weighted) aggregate productivity in the second period, is the (weighted) aggregate productivity of the formal survivors, formalizers, and the new entrants. By subtracting Φ_1 from Φ_2 , it is possible to obtain the productivity change between the first and second period, $\Delta \Phi$, and the contribution of each group.

$$\begin{aligned} \Delta \Phi = & (\Phi_{SF2} - \Phi_{SF1}) + s_{EI2}(\Phi_{EI2} - \Phi_{SF2}) + s_{EN2}(\Phi_{EN2} - \Phi_{SF2}) \\ & + s_{XI1}(\Phi_{SF1} - \Phi_{XI1}) + s_{XT1}(\Phi_{SF1} - \Phi_{XT1}) \end{aligned}$$

The contribution of each group to productivity growth is positive if the productivity among formal survivors increases between the two periods (first term), if informal entrants are more productive than formal survivors in the second period (second term), if new entrants are more productive than formal survivors in the second period (third term), if informal exiters are less productive than formal survivors in the first period (fourth term), or if complete exiters are less productive than formal survivors (fifth term). The contribution of each group of entrants and exiters to aggregate productivity is weighted with the respective group's aggregate market share.¹⁸ For the informal sector, the decomposition can be constructed analogously.

Using this extended decomposition framework, we explore the productivity effects of movements between and complete exits from and new entries into the formal and informal sectors. Beginning with the formal sector, the upper panel of Table 7 shows productivity and market shares of survivors and both types of exiters—informalizers and complete exiters—for the periods to $T=1-2$, $T=2-3$, etc. (i.e. 2005–13). The lower panel then presents TFP and market shares of survivors and both types of entrants—formalizers and new entrants—with reference to $T=2-3$, $T=3-4$, etc. (i.e. 2007–15). Note that the productivity of a group may differ for a given year. For instance, for formal survivors the weighted productivity in 2007 of the first period ($T=1$) differs from its weighted productivity in 2007 of the second period ($T=2$) because formal firms that survived between 2005 and 2007 are not necessarily the same as those that survived between 2007 and 2009. Note that unlike the original MP decomposition, which uses the initial year as the reference for subsequent productivity changes, we are interested in biannual changes between two survey periods.¹⁹

¹⁸ In the first period $s_{SF1} + s_{XI1} + s_{XT1} = 1$ and in the second period $s_{SF2} + s_{EI1} + s_{EN2} = 1$.

¹⁹ The aggregate market share of a group is a function of the number of firms in that group. For instance, if we had used 2005 as the reference year throughout the analyses, the markets share of entrants and exiters relative to survivors would increase over time. This is, in fact, what Melitz and Polanec (2015) do in their original decomposition (see Table 4 of their paper).

Table 7: TFP and market shares of survivors, exiters, and entrants in the formal sector

T=1	TFP			Market share		
	Survivors	Informalizers	Complete exiters	Survivors	Informalizers	Complete exiters
2005	8.60	8.66	8.67	0.92	0.02	0.06
2007	8.85	8.58	9.04	0.80	0.01	0.19
2009	8.90	8.81	8.94	0.78	0.01	0.21
2011	8.97	8.67	9.17	0.79	0.00	0.21
2013	9.09	8.86	9.21	0.81	0.01	0.19

T=2	TFP			Market share		
	Survivors	Formalizers	New entrants	Survivors	Formalizers	New entrants
2007	8.89	8.70	8.91	0.95	0.04	0.02
2009	8.91	8.83	8.89	0.92	0.04	0.05
2011	9.01	8.93	9.20	0.93	0.04	0.03
2013	9.12	9.04	8.90	0.95	0.03	0.02
2015	9.20	8.91	9.21	0.94	0.05	0.02

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

In line with the high rate of exit of formal firms reported in Table 7, the upper panel shows that these exiting firms typically account for about 20 per cent of formal firms' value added. This compares with only 1 or 2 per cent for exiters to informality, which tend to be less productive than formal survivors, and leaves about 80 per cent of the market for surviving firms. Formal firms that become informal are less productive than survivors, but complete exit is not associated with lower productivity. The latter finding merits further investigation, but is not our main focus here.

The lower panel of Table 7 is more interesting, as it shows the characteristics of informal entrants into the formal sector. These formalizers account for 3–4 per cent of formal value added in all but the last period where it rises to 5 per cent. The market share in the last period remains relatively low despite the stark increase in formalization because (previously) informal firms' value added is so low. Looking at TFP, we find that formalizers are not only smaller but also moderately less productive than formal incumbents: In 2007, TFP is 19 per cent lower for formalizers than for formal incumbents and remains at 8 per cent in 2009, 2011, and 2013. With the acceleration in formalization between 2013 and 2015, this difference jumps to more than 30 per cent in 2015. This compares with relatively small TFP differences between (direct) entrants into the formal sector and formal survivors. These new entrants, which account for 2–3 per cent of value added (2007–09 being an exception) are typically as productive as or even more productive than surviving firms in the formal sector. These formal sector dynamics are confirmed when we take labour productivity instead of TFP as our measure of productivity (see Appendix Table A4).

Reproducing Table 7 for the informal sectors mirrors the above patterns, but with reference to informal sector survivors, exiters, and entrants (see Appendix Table A5). While the firms that formalize between 2005 and 2011—typically accounting for about 20–30 per cent of the market—have higher TFP than informal survivors (up to about 30 per cent in 2009), this is not the case in the two last periods. TFP of the many transitioning informal firms, which account for 80 per cent of the market in 2013, is 11 per cent lower than for informal survivors. This confirms our conjecture that it is not only the least productive firms that remain informal. These informal survivors are joined by informalizers that command a relatively high market share that increases to more than 25 per cent in the last period.

The productivity differences between the movers, entrants, and exiters and the market shares drive the respective groups' contribution to productivity growth. Accordingly, Table 8 illustrates that the productivity growth in the formal sector mainly comes from formal survivors. TFP growth

rates of survivors are 10 per cent or higher (over two years) for all periods except 2007–09, the period of the global financial crisis and its repercussions. Because of their considerable market share and their relatively high TFP, complete formal exiters contribute negatively to aggregate TFP growth—up to –4.5 per cent between 2011 and 2013. The contribution of informalizers and new entrants is negligible. In the former case, this is because of a very low market share; in the latter, it is because new entrants resemble survivors in productivity. The key finding of Table 8, however, is the changing contribution of formalizers over time. During the process of continuous formalization until 2013 previously informal sector firms only put a negligible break on formal sector growth—less than 0.5 per cent over two years between 2007 and 2013. Only in the last period, formalizing firms from the informal sector drag down formal sector TFP by –1.3 per cent. For labour productivity (reported in Appendix Table A6), this negative effect of formalizers is stronger. So when looking only at the formal sector, entrants from the informal sector decrease average labour productivity by more than 5 per cent.

Table 8: Contributions to (log) TFP growth in the formal sector

Change	Entrants		Exiters		Formal survivors	Total growth
	Formalizers	New	Informalizers	Complete		
2005–07	–0.008	0.000	–0.001	–0.004	0.291	0.28
2007–09	–0.003	–0.001	0.003	–0.036	0.061	0.02
2009–11	–0.003	0.005	0.001	–0.008	0.106	0.10
2011–13	–0.003	–0.004	0.000	–0.045	0.156	0.11
2013–15	–0.013	0.000	0.002	–0.022	0.104	0.07

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Table 9 shows the growth contributions of the different groups in the informal sector. Again, survivors' performance drives average TFP growth. Growth rates are high among informal survivors—probably contrary to expectations. There is a large drop in TFP between 2007 and 2009, suggesting that informal firms were hit much harder by the economic crisis of these years. The few informal survivors of the last period exhibit very high TFP growth of well above 50 per cent. New entrants contribute positively but very moderately to TFP growth while complete exit does not necessarily make a positive contribution to TFP growth, similar to what we have seen in the formal sector. In contrast, the transitions from and into formality affect average informal TFP growth considerably. With the exception of the first period, informalizers contribute negatively to TFP growth and increasingly so. In contrast, the firms that exit informality (i.e. the formalizers) make an important positive contribution to informal TFP growth after 2011 because they are less productive than informal survivors when they leave the informal sector. This is different in earlier years when relatively more productive firms formalized, thereby decreasing average informal sector TFP growth. Our interpretation of this pattern is a 'regime switch' from endogenous to exogenous policy-driven selection that alters the characteristics of formalizers—in line with the descriptive statistics provided in Table 5.

Table 9: Contributions to (log) TFP growth in the informal sector

Change	Entrants		Exiters		Informal survivors	Total (log) growth
	Informalizers	New	Formalizers	Complete		
2005–07	0.046	0.003	–0.037	–0.009	0.347	0.35
2007–09	–0.004	0.004	–0.020	0.016	–0.056	–0.06
2009–11	–0.013	0.015	–0.086	0.004	0.173	0.09
2011–13	–0.057	0.009	0.067	–0.008	0.100	0.11
2013–15	–0.116	0.007	0.081	–0.026	0.539	0.48

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Finally, we return to the insight of the OP decomposition that the weighted productivity of any group can be written as the sum of its unweighted productivity and the covariance between market shares and productivity. Table 10 that depicts the contribution of these two terms to weighted productivity for formal survivors, formalizers, informalizers, and informal survivors requires some explanation.²⁰ The columns 'Unw.' (i.e. unweighed) and 'Cov.' (i.e. covariance) add to within-group value-added weighted TFP growth. For example, it is 0.29 in log TFP growth for formal survivors, of which 0.23 comes from growth in unweighted TFP and 0.06 from changes in TFP dispersion among surviving formal firms. This group's growth rate is then weighted with the value-added share (column 'Share') of survivors in total value added of all firms, which allows us to decompose aggregate TFP growth—of SME manufacturing firms—into the contributions of each of the groups. Further, for presentational reasons, we only add one column for the net growth effect of entry and exit (for entrants and exiters we cannot observe changes in the distribution of TFP).²¹

Again, we can see that formal survivors drive overall TFP growth and that this growth partly stems from improved resource allocation among surviving firms. More productive formal survivors thus gain market shares, an effect that is particularly pronounced between 2011 and 2013. Because of the decreasing share of informal production, the contribution of informal firms to overall growth declines, but there are interesting firm dynamics at work. During the crisis between 2007 and 2009, more productive informal firms lose market share, which is possibly related to the fact that export-oriented (and often more productive) firms were more heavily affected. In contrast, resource allocation improvements explain all of the high TFP growth among informal survivors in the last period. Because of the small sample size in 2013 we do not want to overemphasize this result.²²

²⁰ The detailed decomposition results for both T=1 and T=2 are provided in Appendix Table A7 for TFP and Appendix Table A8 for labour productivity.

²¹ As the aggregate market shares of the respective groups also change between the years, this would add another growth component that we ignore by using each group's average market share over the two-year periods. In addition, sample replacements affect the reported figures although the effects should be minor.

²² We can count only on 25 informal survivors in 2013, but replacements raise this number to 187 in 2015, so we are comparing two potentially fairly different sets of firms (see Appendix Table A7). There is no other group where the group composition changes so dramatically.

Table 10: Contributions to total (log) TFP growth

Period	Formal survivors				Formalizers				Informalizer				Informal survivors				Entry and exit		Overall		
	Unw.	Cov.	Share	Contr.	Unw.	Cov.	Share	Contr.	Unw.	Cov.	Share	Contr.	Unw.	Cov.	Share	Contr.	Share	Contr.	Unw.	Cov.	Total
2005–07	0.23	0.06	0.84	0.24	0.81	0.17	0.03	0.03	–0.38	–0.11	0.02	–0.01	0.18	0.16	0.08	0.03	0.04	0.01	0.24	0.04	0.28
2007–09	0.04	0.02	0.79	0.05	0.60	0.23	0.03	0.02	–0.50	–0.22	0.01	–0.01	0.04	–0.09	0.06	–0.00	0.12	0.01	0.09	–0.03	0.06
2009–11	0.10	0.01	0.80	0.09	0.68	0.12	0.03	0.02	–0.54	–0.31	0.01	–0.01	0.15	0.02	0.04	0.01	0.12	0.04	0.17	0.00	0.17
2011–13	0.06	0.10	0.82	0.13	1.03	0.19	0.02	0.03	–1.12	–0.78	0.00	–0.00	0.10	0.00	0.04	0.00	0.12	–0.02	0.07	–0.01	0.06
2013–15	0.09	0.02	0.84	0.09	0.53	0.32	0.04	0.04	–0.49	–0.13	0.01	–0.00	–0.02	0.55	0.01	0.00	0.10	–0.00	0.14	–0.01	0.13

Note: Unw., unweighted; Cov., covariance; Contr, contribution. Market shares are averages over the respective periods. For entry and exit, the market share is the average of exiters in period 1 and entrants in period 2. This explains why the ‘shares’ do not exactly add up to 1.

Source: authors’ calculation based on UNU-WIDER’s Viet Nam SME database (see CIEM et al. 2005–15).

One of the key results of our empirical analysis is what happens to formalizers. Unweighted TFP growth of formalizers is very strong, typically more than 80 per cent over each two-year period. Only in the last period is it somewhat lower at about 70 per cent. This brings them closer to the average TFP levels of formal survivors, in particular in the years up to 2013 (see Appendix Table A7). In addition, among formalizing firms, more productive firms expand much more quickly than less productive ones. The covariance term adds another 12–32 log points to the TFP growth rate of formalizers. Formalizers matter for aggregate manufacturing SME productivity: they contribute between 2 and 4 percentage points to overall (two-year) TFP growth.

The much smaller group of informalizers has little aggregate impact. Yet, the dynamics are also noteworthy. In this group, not only is (unweighted) average TFP growth negative but also the contribution of changes in the covariance term is negative; that is, among informalizing firms relatively less productive firms gain market share. The net contribution of entry and exit (including to and from both formal and informal sectors) is relatively small. This is because some of the larger effects of entry and exit tend to cancel out each other. We have already shown that exiters are not among the least productive firms, contrary to expectations. The covariance terms tend to be much higher for these two groups than for the other groups; that is, entering and exiting firms with higher TFP tend to have higher market shares.

We also conduct the decomposition analysis for labour productivity and many of the results point to a similar direction. One important exception is the slight decrease of (weighted) labour productivity among formalizers in the last period that accompanies the strong TFP increase (see Appendix Table A8). In addition, in contrast to TFP, the covariance term decreases for labour productivity, implying that the firms with TFP gains are relatively labour-intensive. Further examination of these opposite movements of TFP and labour productivity shows that formalizers between 2013 and 2015 do not expand their workforce, which implies that formalizers adjust their capital stocks (downwards). In earlier periods, formalizers do expand their workforce moderately.

6 Conclusion

This study has examined the productivity dynamics of manufacturing SMEs in Viet Nam. Over the ten years (2005–15) covered by the unique panel dataset that we use for our analysis, the country's manufacturing sector has rapidly formalized, with an acceleration of formalization likely to be triggered by policy reforms that reinforce compliance with registration requirements, including higher fines. Our descriptive analyses based on productivity decompositions proposed by Olley and Pakes (1996) and Melitz and Polanec (2015) disentangle misallocation among and aggregate productivity contributions of formal, informal, formalizing, and informalizing firms as well as of entry and exit.

The static decompositions confirm a substantial productivity gap between formal and informal firms, which is partly due to higher misallocation in the informal sector. Our findings fit well with a view of firms weighing the costs and benefits of formalization, with policy reforms apparently raising considerably the costs of remaining informal. This leads to a 'regime change': between 2005 and 2013, there is a continuous process of formalization with formalization rates typically about 15–20 per cent (over two-year periods). In this process, more productive informal firms select into formality compared with informal survivors. The formalizers are a dynamic group of firms. It 'pays' for them to formalize, as they see very high TFP gains, in particular for high productivity firms. These firms gain market share, implying that there are aggregate allocative gains among formalizers.

Between 2013 and 2015, the formalization rate jumps to 80 per cent (of informal firms in 2015), including now quite a number of less productive firms. Although this regime change affects formal sector TFP growth negatively, there are aggregate productivity gains. This is because the formalization-induced dynamics—high TFP growth and improved resource allocation—remain intact. In this process, there also appear to be some adjustment costs for formalizers, which on average reduce their capital stocks. These results suggest that Viet Nam’s informal manufacturing firms—or the large majority of them—cannot be seen as subsistence activities. Despite continuous formalization, informal survivors are not merely left-behind low-productivity firms.

There are a couple of additional insights from our dynamic decomposition. Although our focus is on formalizing firms, the increasing importance of formal manufacturing SMEs implies that their productivity performance is the dominant driver of overall productivity. Entry and exit are very important in terms of market share and employment. Yet, the contribution to productivity growth is modest because exiters are not (much) less productive than survivors. Opposite to what one might expect churning is not higher in the informal sector. In fact, exit rates are even higher in the formal sector. This, again, speaks against informal firms being stagnant (and volatile) subsistence-type activities.

Overall, our analysis is in line with the evidence that suggests positive impacts of formalization on firm performance in Viet Nam. Formalization appears to help previously informal firms to thrive and this effect seems stronger for more productive firms. The results indicate that formalization is associated with adjustments: higher TFP, but also adjustments to capital stocks, which may also impose costs on firm owners and employees.

Our study certainly warns against generalizations across time, place, and policies as well as across industries. It is not obvious that the findings from Viet Nam with its relatively large formal sector to start with—at least in manufacturing—generalize to other country context with much higher levels of informality, for example, in Latin America. Formality and formalization have to do with how citizens—here, of course, particularly firm owners—perceive the state. In particular, the willingness to pay taxes—an important aspect of business registration—depends on whether taxes are merely perceived as a fee without return. It may well be that the regulatory changes, enforcement, and higher fines may not induce formalization similar to what we have seen in Viet Nam. In particular, policies may trigger context-dependent heterogeneous selection processes that eventually result in very different impacts on firm performance and aggregate productivity. Generalizations may also be limited because our analysis is confined to the manufacturing sector. More persistent forms of low-productivity informal self-employment are likely to be found in the service sector.

In the context that we have investigated, our analysis suggests that increasing formality through tighter rules and enforcement has been a good policy choice. Yet, it is probably too early for a final judgement of the impacts of this formalization wave on firm performance. Our results are indicative of rather positive formalization impacts, but future research should, for example, look into exit rates among formalizing firms. Furthermore, formalization has benefits beyond firm performance that we have ignored; for example, in terms of the formalization of labour, access to social protection, and the mobilization of tax revenue. Finally, our study suggests careful analysis of formal and informal firm dynamics when judging aggregate productivity performance of a developing economy. When previously informal firms formalize and become part of the statistically measured economy they may exert downward pressure on average (formal) productivity, which may lead to misguided interpretations of economic performance.

References

- Akerberg, D., C. Lanier Benkard, S. Berry, and A. Pakes (2007). 'Econometric Tools for Analyzing Market Outcomes'. In J.J. Heckman and E.E. Leamer (eds), *Handbook of Econometrics* (Volume 6, Chapter 63, pp. 4171–276). [https://doi.org/10.1016/S1573-4412\(07\)06063-1](https://doi.org/10.1016/S1573-4412(07)06063-1)
- Benhassine, N., D. McKenzie, V. Pouliquen, and M. Santini (2018). 'Does Inducing Informal Firms to Formalize Make Sense? Experimental Evidence from Benin'. *Journal of Public Economics*, 157: 1–14. <https://doi.org/10.1016/j.jpubeco.2017.11.004>
- Blattman, C., E.P. Green, J. Jamison, M.C. Lehmann, and J. Annan (2016). 'The Returns to Microenterprise Support among the Ultrapoorest: A Field Experiment in Postwar Uganda'. *American Economic Journal: Applied Economics*, 8: 35–64. <https://doi.org/10.1257/app.20150023>
- Bruhn, M., and D. McKenzie (2014). 'Entry Regulation and the Formalization of Microenterprises in Developing Countries'. *The World Bank Research Observer*, 29: 186–201. <https://doi.org/10.1093/wbro/lku002>
- Busso, M., M.V. Fazio, and S.L. Algazi (2012). '(In)Formal and (Un)Productive: The Productivity Costs of Excessive Informality in Mexico'. Inter-American Development Bank Working Paper IDB-WP-341. Washington, DC: Inter-American Development Bank (IDB). <https://doi.org/10.2139/ssrn.2207240>
- CIEM, ILSSA, UCPH, and UNU-WIDER (2005–15). Viet Nam SME Survey. Available at: <https://www.wider.unu.edu/project/small-and-medium-enterprise-sme-survey> (accessed 30 October 2020).
- Cling, J.-P., M. Razafindrakoto, and F. 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: 632–52. <https://doi.org/10.1080/13547860.2012.724553>
- Cusolito, A.P., and W.F. Maloney (2019). *Productivity Revisited: Shifting Paradigms in Analysis and Policy*. Washington, DC: World Bank. <https://doi.org/10.1596/978-1-4648-1334-4>
- David, J.M., and V. Venkateswaran (2019). 'The Sources of Capital Misallocation'. *American Economic Review*, 109: 2531–67. <https://doi.org/10.1257/aer.20180336>
- David, J.M., V. Venkateswaran, A.P. Cusolito, and T. Didier (2019). 'Capital Allocation in Developing Countries'. Policy Research Working Papers. Washington, DC: World Bank. <https://doi.org/10.1596/1813-9450-9031>
- De Loecker, J., P.K. Goldberg, A.K. Khandelwal, N. Pavcnik (2016). 'Prices, Markups, and Trade Reform'. *Econometrica*, 84: 445–510. <https://doi.org/10.3982/ECTA11042>
- de Mel, S., D. McKenzie, and C. Woodruff (2008). 'Returns to Capital in Microenterprises: Evidence from a Field Experiment'. *The Quarterly Journal of Economics*, 123: 1329–72. <https://doi.org/10.1162/qjec.2008.123.4.1329>
- de Mel, S., D. McKenzie, and C. Woodruff (2013). 'The Demand for, and Consequences of, Formalization among Informal Firms in Sri Lanka'. *American Economic Journal: Applied Economics*, 5: 122–50. <https://doi.org/10.1257/app.5.2.122>
- Demenet, A., M. Razafindrakoto, and F. Roubaud (2016). 'Do Informal Businesses Gain from Registration and How? Panel Data Evidence from Vietnam'. *World Development*, 84: 326–41. <https://doi.org/10.1016/j.worlddev.2015.09.002>
- DERG (2005–10). *Vietnam Data: Enterprise and Household Data*. Copenhagen: University of Copenhagen. Available at: <https://www.econ.ku.dk/derg/links/vietnam> (accessed 30 October 2020).
- Du, J., X. Liu, and Y. Zhou (2014). 'State Advances and Private Retreats?—Evidence of Aggregate Productivity Decomposition in China'. *China Economic Review*, 31: 459–74. <https://doi.org/10.1016/j.chieco.2014.03.002>

- Fafchamps, M., D. McKenzie, S. Quinn, and C. Woodruff (2014). 'Microenterprise Growth and the Flypaper Effect: Evidence from a Randomized Experiment in Ghana'. *Journal of Development Economics*, 106: 211–26. <https://doi.org/10.1016/j.jdeveco.2013.09.010>
- Foster, L., J. Haltiwanger, and C. Syverson (2008). 'Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?'. *The American Economic Review*, 98: 394–425. <https://doi.org/10.1257/aer.98.1.394>
- Gollin, D., and C.R. Udry (2019). 'Heterogeneity, Measurement Error and Misallocation: Evidence from African Agriculture'. NBER Working Paper 25440. Cambridge, MA: National Bureau of Economic Research (NBER). <https://doi.org/10.3386/w25440>
- Grimm, M., P. Knorringa, and J. Lay (2012). 'Constrained Gazelles: High Potentials in West Africa's Informal Economy'. *World Development*, 40: 1352–68. <https://doi.org/10.1016/j.worlddev.2012.03.009>
- Grimm, M., J. Krüger, and J. Lay (2011). 'Barriers to Entry and Returns to Capital in Informal Activities: Evidence from Sub-Saharan Africa'. *Review of Income and Wealth*, 57: 27–53. <https://doi.org/10.1111/j.1475-4991.2011.00453.x>
- Government of Viet Nam (2006). Decree No. 88/2006/ND-CP. Available at: <https://vanbanphapluat.co/decreed-no-88-2006-nd-cp-of-august-29-2006-on-business-registration> (accessed October 2020).
- Government of Viet Nam (2013). Decree No. 185/2013/ND-CP. Available at: <https://vietnamlawenglish.blogspot.com/2013/11/vietnam-decree-no-1852013nd-cp-on.html> (accessed October 2020).
- GSO (2004). 'Results of Establishment Census of Vietnam 2002'. Hanoi: Statistical Publishing House.
- Hsieh, C.-T., and P.J. Klenow (2009). 'Misallocation and Manufacturing TFP in China and India'. *The Quarterly Journal of Economics*, 124: 1403–48. <https://doi.org/10.1162/qjec.2009.124.4.1403>
- Kinghan, C., C. Newman, and C.M. O'Toole (2018). 'Capital Allocation, Credit Access, and Firm Growth in Vietnam'. WIDER Working Paper 67/2018. Helsinki: UNU-WIDER. <https://doi.org/10.35188/UNU-WIDER/2018/509-1>
- La Porta, R., and A. Shleifer (2014). 'Informality and Development'. *Journal of Economic Perspectives*, 28: 109–26. <https://doi.org/10.1257/jep.28.3.109>
- Levinsohn, J., and A. Petrin (2003). 'Estimating Production Functions Using Inputs to Control for Unobservables'. *Review of Economic Studies*, 70(2): 317–41. <https://doi.org/10.1111/1467-937X.00246>
- Levy, S., and L.F. López-Calva (2018). 'Persistent Misallocation and the Returns to Education in Mexico'. *The World Bank Economic Review*, 34(2): 284–311. <https://doi.org/10.1093/wber/lhy017>
- McCaig, B., and J. Nanowski (2019). 'Business Formalisation in Vietnam'. *The Journal of Development Studies*, 55: 805–21. <https://doi.org/10.1080/00220388.2018.1475646>
- McCaig, B., and N. Pavcnik (2017). 'Out with the Old and Unproductive, In with the New and Similarly Unproductive: Microenterprise Dynamics in a Growing Low-Income Economy'. GLM|LIC Working Paper 23. Bonn: Institute of Labor Economics. Available at: <https://g2lm-lic.iza.org/publications/wp/wp23/> (accessed November 2020).
- McCaig, B., and N. Pavcnik (2018). 'Export Markets and Labor Allocation in a Low-Income Country'. *American Economic Review*, 108: 1899–941. <https://doi.org/10.1257/aer.20141096>
- McKenzie, D., and Y.S. Sakho (2010). 'Does It Pay Firms to Register for Taxes? The Impact of Formality on Firm Profitability'. *Journal of Development Economics*, 91: 15–24. <https://doi.org/10.1016/j.jdeveco.2009.02.003>
- McKenzie, D., and C. Woodruff (2008). 'Experimental Evidence on Returns to Capital and Access to Finance in Mexico'. *The World Bank Economic Review*, 22(3): 457–82. <https://doi.org/10.1093/wber/lhn017>

- Melitz, M.J., and S. Polanec (2015). ‘Dynamic Olley–Pakes Productivity Decomposition with Entry and Exit’. *The RAND Journal of Economics*, 46: 362–75. <https://doi.org/10.1111/1756-2171.12088>
- Olley, G.S., and A. Pakes (1996). ‘The Dynamics of Productivity in the Telecommunications Equipment Industry’. *Econometrica*, 64: 1263. <https://doi.org/10.2307/2171831>
- Perry, G.E., O. Arias, P. Fajnzylber, W.F. Maloney, A. Mason, and J. Saavedra-Chanduvi (2007). ‘Informality: Exit and Exclusion’. Washington, DC: The World Bank. <https://doi.org/10.1596/978-0-8213-7092-6>
- Rand, J., and F. Tarp (2020). *Micro, Small, and Medium Enterprises in Vietnam*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198851189.003.0001>
- Rand, J., and N. Torm (2012). ‘The Benefits of Formalization: Evidence from Vietnamese Manufacturing SMEs’. *World Development*, 40: 983–98. <https://doi.org/10.1016/j.worlddev.2011.09.004>
- Restuccia, D., and R. Rogerson (2008). ‘Policy Distortions and Aggregate Productivity with Heterogeneous Establishments’. *Review of Economic Dynamics*, 11: 707–20. <https://doi.org/10.1016/j.red.2008.05.002>
- Restuccia, D., and R. Rogerson (2017). ‘The Causes and Costs of Misallocation’. *Journal of Economic Perspectives*, 31: 151–74. <https://doi.org/10.1257/jep.31.3.151>
- Restuccia, D., and R. Santaaulalia-Llopis (2017). ‘Land Misallocation and Productivity’. NBER Working Paper 23128. Cambridge, MA: National Bureau of Economic Research (NBER). <https://doi.org/10.3386/w23128>
- Syverson, C. (2004). ‘Market Structure and Productivity: A Concrete Example’. *Journal of Political Economy*, 112: 1181–222. <https://doi.org/10.1086/424743>
- Ulyssea, G. (2018). ‘Firms, Informality, and Development: Theory and Evidence from Brazil’. *American Economic Review*, 108: 2015–47. <https://doi.org/10.1257/aer.20141745>
- Williams, C.C., M.S. Shahid, A. Martínez (2016). ‘Determinants of the Level of Informality of Informal Micro-Enterprises: Some Evidence from the City of Lahore, Pakistan’. *World Development*, 84: 312–25. <https://doi.org/10.1016/j.worlddev.2015.09.003>

Appendix A: Additional tables

Table A1: Sectoral distribution of formal and informal firms

Two-digit ISIC sector	Formal firm (firm has a tax code number)					
	No		Yes		Total	
	No.	Col. %	No.	Col. %	No.	Col. %
Food and beverages	737	41.9	635	23.3	1,372	30.6
Tobacco	6	0.3	4	0.1	10	0.2
Textiles	62	3.5	105	3.9	167	3.7
Apparel	38	2.2	189	6.9	227	5.1
Leather	22	1.3	78	2.9	100	2.2
Wood	275	15.6	195	7.2	470	10.5
Paper	15	0.9	111	4.1	126	2.8
Publishing and printing	9	0.5	123	4.5	132	2.9
Refined petroleum, etc.	6	0.3	7	0.3	13	0.3
Chemical products, etc.	9	0.5	95	3.5	104	2.3
Rubber	14	0.8	199	7.3	213	4.7
Non-metallic mineral products	107	6.1	139	5.1	246	5.5
Basic metals	12	0.7	42	1.5	54	1.2
Fabricated metal products	266	15.1	471	17.3	737	16.4
Electronic machinery, etc.	10	0.6	121	4.4	131	2.9
Motor vehicles, etc.	3	0.2	29	1.1	32	0.7
Other transport equipment	1	0.1	9	0.3	10	0.2
Furniture, etc.	155	8.8	146	5.4	301	6.7
Recycling, etc.	6	0.3	8	0.3	14	0.3
Manufacture and distribution of electricity, etc.	5	0.3	21	0.8	26	0.6
Total	1,758	100.0	2,727	100.0	4,485	100.0

Note: ISIC, International Standard Industrial Classification. A firm's sector is the one that it operated in when surveyed for the first time.

Source: authors' calculation based on UNU-WIDER's Viet Nam small and medium-sized enterprise (SME) database.

Table A2: Coefficient and productivity estimates

Sector	OLS estimates			FE estimates			ACF estimates		
	Labour	Capital	TFP	Labour	Capital	TFP	Labour	Capital	TFP
Informal sector	1.07	0.15	7.68	0.67	0.06	9.22	0.98	0.16	7.70
Formal sector	1.04	0.13	8.39	0.74	0.09	9.54	0.99	0.13	8.50

Note: OLS, ordinary least-square; FE, fixed-effect; ACF, Akerberg et al. (2007); TFP, total factor productivity. The table depicts value added and capital elasticities that are estimated via OLS, FE, and ACF routines. TFP is subsequently computed as the residual of the previous production function estimation.

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Table A3: Biannual transitions, 2005–15

					Status									
2007					2009					2011				
Informal	Formal	Exit	Total		Informal	Formal	Exit	Total		Informal	Formal	Exit	Total	
Status 2005					Status 2007					Status 2009				
Informal	807	158	70	1,035	Informal	694	152	153	999	Informal	549	160	167	876
Formal	86	1,088	94	1,268	Formal	57	1,090	300	1,447	Formal	31	1,171	370	1,572
Not in sample	110	210	0	320	Not in sample	145	364	0	509	Not in sample	122	328	0	450
Total	1,003	1,456	164	2,623	Total	896	1,606	453	2,955	Total	702	1,659	537	2,898
Not in sample 2005					Not in sample 2007					Not in sample 2009				
Informal at start-up	88	19	0	107	Informal at start-up	121	31	0	152	Informal at start-up	97	53	0	150
Formal at start-up	11	147	0	158	Formal at start-up	6	240	0	246	Formal at start-up	5	199	0	204
New entrant	11	44	0	55	New entrant	18	93	0	111	New entrant	20	76	0	96
Total	110	210	0	320	Total	145	364	0	509	Total	122	328	0	450
					Status									
2013					2015									
Informal	Formal	Exit	Total		Informal	Formal	Exit	Total						
Status 2011					Status 2013									
Informal	543	54	102	699	Informal	26	568	113	707					
Formal	4	1,303	319	1,626	Formal	35	1,346	304	1,685					
Not in sample	160	328	0	488	Not in sample	202	316	0	518					
Total	707	1,685	421	2,813	Total	263	2,230	417	2,910					
Not in sample 2011					Not in sample 2013									
Informal at start-up	130	36	0	166	Informal at start-up	167	48	0	215					
Formal at start-up	9	252	0	261	Formal at start-up	3	221	0	224					
New entrant	21	40	0	61	New entrant	32	47	0	79					
Total	160	328	0	488	Total	202	316	0	518					

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Table A4: Labour productivity and market shares of survivors, exiters, and entrants in the formal sector

T=1	Labour productivity			Market share		
	Survivors	Exiters to informality	Complete exiters	Survivors	Exiters to informality	Complete exiters
2005	9.88	9.53	9.86	0.92	0.03	0.05
2007	10.19	9.68	10.28	0.82	0.02	0.16
2009	10.28	10.09	10.27	0.80	0.01	0.19
2011	10.44	9.93	10.53	0.81	0.00	0.19
2013	10.42	10.14	10.47	0.82	0.01	0.17
T=2	Labour productivity			Market share		
	Survivors	Informal entrants	New entrants	Survivors	Informal entrants	New entrants
2007	10.23	9.86	9.92	0.92	0.05	0.02
2009	10.29	10.06	10.29	0.91	0.05	0.04
2011	10.47	10.25	10.28	0.92	0.05	0.03
2013	10.43	10.30	10.40	0.95	0.03	0.02
2015	10.50	9.82	10.42	0.90	0.08	0.02

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Table A5: TFP and market shares of survivors, exiters, and entrants in the informal sector

T=1	TFP			Market share		
	Survivors	Exiters to formality	Complete exiters	Survivors	Exiters to formality	Complete exiters
2005	7.56	7.72	7.75	0.72	0.23	0.05
2007	7.95	8.03	7.83	0.63	0.25	0.13
2009	7.82	8.09	7.80	0.47	0.31	0.22
2011	8.05	7.73	8.11	0.66	0.21	0.13
2013	8.16	8.05	8.35	0.06	0.80	0.14
T=2	TFP			Market share		
	Survivors	Entrants from formality	New entrants	Survivors	Entrants from formality	New entrants
2007	7.91	8.22	8.20	0.84	0.15	0.01
2009	7.90	7.87	8.07	0.85	0.12	0.02
2011	7.99	7.85	8.44	0.88	0.09	0.03
2013	8.15	6.77	8.48	0.93	0.04	0.03
2015	8.69	8.24	8.78	0.66	0.26	0.08

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Table A6: Contributions to (log) labour productivity growth in the formal sector

Change	Entrants		Exiters		Formal survivors	Total growth
	From informality	New	To informality	Complete		
2005–07	–0.020	–0.007	0.009	0.001	0.348	0.33
2007–09	–0.012	0.000	0.009	–0.014	0.093	0.08
2009–11	–0.011	–0.006	0.002	0.002	0.190	0.18
2011–13	–0.004	–0.001	0.000	–0.017	–0.003	–0.02
2013–15	–0.052	–0.001	0.003	–0.008	0.075	0.02

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Table A7: Contributions to overall TFP

	Formal survivors					Formalizers					Informalizer								
	Wei.	Unw.	Cov.	Share	Obs.	Wei.	Unw.	Cov.	Share	Obs.	Wei.	Unw.	Cov.	Share	Obs.				
Period 2005–07																			
2005	8.60	8.40	0.20	0.81	1,064	7.72	7.71	0.00	0.03	149	8.66	8.37	0.27	0.02	83				
2007	8.89	8.63	0.27	0.86	1,203	8.70	8.52	0.18	0.03	173	8.22	7.99	0.17	0.01	90				
Period 2007–09																			
2007	8.85	8.63	0.22	0.72	1,070	8.03	8.05	−0.06	0.02	147	8.58	8.46	0.12	0.01	57				
2009	8.91	8.67	0.24	0.85	1,227	8.83	8.66	0.17	0.03	171	7.87	7.96	−0.10	0.01	61				
Period 2009–11																			
2009	8.90	8.68	0.23	0.73	1,100	8.09	8.08	0.04	0.02	149	8.81	8.53	0.28	0.01	30				
2011	9.01	8.77	0.24	0.88	1,338	8.93	8.75	0.15	0.04	202	7.85	8.00	−0.04	0.01	29				
Period 2011–13																			
2011	8.97	8.77	0.20	0.74	1,271	7.73	7.78	0.03	0.01	49	8.67	8.70	−0.03	0.00	4				
2013	9.12	8.82	0.30	0.90	1,521	9.04	8.80	0.22	0.03	86	6.77	7.58	−0.81	0.00	12				
Period 2013–15																			
2013	9.09	8.80	0.30	0.76	1,315	8.05	7.96	0.09	0.04	556	8.86	8.72	0.14	0.01	35				
2015	9.20	8.89	0.32	0.92	1,529	8.91	8.50	0.41	0.05	606	8.24	8.23	0.01	0.01	38				
	Informal survivors					Exiters					Entrants								
	Wei.	Unw.	Cov.	Share	Obs.	Wei.	Unw.	Cov.	Share	Obs.	Wei.	Unw.	Cov.	Share	Obs.				
Period 2005–07																			
2005	7.56	7.51	0.06	0.09	786	8.57	8.01	0.56	0.06	157	8.87	8.56	0.31	0.02	54				
2007	7.91	7.69	0.21	0.08	871	8.96	8.25	0.72	0.19	435									
Period 2007–09																			
2007	7.95	7.68	0.26	0.06	682					8.86	8.66	0.22	0.05	105					
2009	7.90	7.72	0.17	0.06	755														
Period 2009–11																			
2009	7.82	7.70	0.11	0.03	520	8.86	8.35	0.50	0.21	520	9.14	8.80	0.35	0.03	95				
2011	7.99	7.85	0.13	0.05	632	9.13	8.58	0.55	0.21	443									
Period 2011–13																			
2011	8.05	7.89	0.15	0.04	529					8.86	8.46	0.40	0.02	60					
2013	8.15	8.00	0.15	0.05	652														
Period 2013–15																			
2013	8.16	8.20	−0.05	0.00	25	9.18	8.70	0.47	0.18	400	9.17	8.67	0.50	0.02	76				
2015	8.69	8.19	0.50	0.01	187														

Note: Wei., weighted; Unw., unweighted; Cov., covariance; Obs., observations.

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Table A8: Contributions to overall labour productivity

	Formal survivors					Formalizers					Informalizer									
	Wei.	Unw.	Cov.	Share	Obs.	Wei.	Unw.	Cov.	Share	Obs.	Wei.	Unw.	Cov.	Share	Obs.					
Period 2005–07																				
2005	9.88	9.89	−0.01	0.81	1,087	9.39	9.45	−0.07	0.03	158	9.53	9.60	−0.07	0.02	86					
2007	10.23	10.20	0.03	0.86	1,235	9.86	9.89	−0.03	0.03	177	9.89	9.83	0.05	0.01	97					
Period 2007–09																				
2007	10.19	10.16	0.03	0.72	1,090	9.69	9.81	−0.12	0.02	152	9.68	9.80	−0.12	0.01	57					
2009	10.29	10.21	0.08	0.85	1,292	10.06	9.93	0.13	0.03	180	9.80	9.86	−0.06	0.01	63					
Period 2009–11																				
2009	10.28	10.17	0.11	0.73	1,162	9.89	9.89	0.00	0.02	157	10.09	9.88	0.21	0.01	31					
2011	10.47	10.34	0.13	0.88	1,370	10.25	10.14	0.11	0.04	213	10.29	10.03	0.26	0.01	36					
Period 2011–13																				
2011	10.44	10.29	0.15	0.74	1,303	10.37	10.29	0.09	0.01	54	9.93	10.03	−0.09	0.00	4					
2013	10.43	10.31	0.12	0.90	1,555	10.30	10.25	0.05	0.03	90	10.12	10.20	−0.08	0.00	13					
Period 2013–15																				
2013	10.42	10.31	0.11	0.76	1,346	9.84	9.66	0.18	0.04	568	10.14	10.08	0.07	0.01	35					
2015	10.50	10.39	0.11	0.92	1,567	9.82	9.70	0.12	0.05	616	10.07	10.07	0.00	0.01	38					
	Informal survivors					Exiters					Entrants									
	Wei.	Unw.	Cov.	Share	Obs.	Wei.	Unw.	Cov.	Share	Obs.	Wei.	Unw.	Cov.	Share	Obs.					
Period 2005–07																				
2005	9.34	9.27	0.07	0.09	807	9.65	9.55	0.10	0.06	165	9.87	9.98	−0.11	0.02	55					
2007	9.50	9.46	0.04	0.08	807	10.14	9.93	0.21	0.19	466										
Period 2007–09											10.28	10.23	0.05	0.05	109					
2007	9.52	9.44	0.09	0.06	694															
2009	9.59	9.45	0.14	0.06	686															
Period 2009–11											10.25	10.17	0.08	0.03	96					
2009	9.50	9.38	0.12	0.03	549	10.20	9.98	0.21	0.21	550										
2011	9.88	9.65	0.23	0.05	549															
Period 2011–13											10.32	10.01	0.32	0.02	61					
2011	9.81	9.63	0.18	0.04	543	10.48	10.21	0.27	0.21	457										
2013	9.80	9.63	0.17	0.05	543															
Period 2013–15											10.33	10.18	0.15	0.02	79					
2013	10.11	10.00	0.11	0.00	26	10.42	10.14	0.28	0.18	417										
2015	10.27	10.25	0.03	0.01	26															

Note: Wei., weighted; Unw., unweighted; Cov., covariance; Obs., observations.

Source: authors' calculation based on UNU-WIDER's Viet Nam SME database (see CIEM et al. 2005–15).

Appendix B: ACF estimation procedure

Our primary measure of productivity, total factor productivity (TFP), is computed as the residual from the estimation of a value added Cobb–Douglas production function.^{*} Hence, the production function (in logs) we take to the data, and estimate is

$$y_{it} = \alpha_l l_{it} + \alpha_k k_{it} + \varphi_{it} + \epsilon_{it}$$

where y_{it} is value added, l_{it} is the number of employees, and k_{it} is the capital stock, all measured for firm i at time t .[†] φ_{it} captures a firm’s productivity at time t and subsumes the constant term, and ϵ_{it} is a standard independent and identically distributed error term.

To obtain a firm’s value added, we subtract from a firm’s output its intermediate inputs, defined as the consumption of commodities by a firm in the manufacturing process. It includes raw material expenses and other indirect costs.[‡] Capital stocks are plants’ reported book values of land, buildings, and machines. Output, intermediate inputs, and capital stocks are deflated by detailed producer price indexes from the General Statistics Office of Viet Nam. Labour is measured using the total number of full-time regular employees at a plant. We prefer this measure to the wage bill because for many firms a considerable fraction of the labour force is unpaid—66 per cent and 27 per cent for formal and informal firms, respectively—and would therefore not be captured by the wage bill.[§]

It is well known that ordinary least squares (OLS) leads to biased estimates of the production function as the inputs are chosen endogenously, and that the bias will be more severe for those inputs that respond quickly to a productivity shock and thus are freely adjustable (here, labour). Hence, to consistently estimate that the parameters of the production function, which, in the case of a Cobb–Douglas production function, constitute the input elasticities, we rely on the proxy estimator introduced by Akerberg et al. (2007) (henceforth, ACF). The ACF estimator corrects for the selection and simultaneity biases while relying on a model of industry dynamics. Notice that the TFP measure we generate following the ACF procedure is based on a production function in revenues.

Crucial for the identification of the production function parameters is the assumption that firms use some dynamic and some variable inputs into production. Typically, capital is assumed to be the dynamic input subject to adjustment costs whereas labour is assumed to be the variable input

^{*} As an additional robustness test, we generate our productivity measure estimating the more flexible translog production function.

[†] In further discussion here, small letters denote logs.

[‡] Indirect costs include, among others, expenses on electricity, fuel, water, rental costs, stationery and office supplies, telephone/internet, transport, insurance, marketing and advertisement, and maintenance and repair.

[§] Using the wage bill has the advantage of capturing marginal products of heterogeneous workers and thus accounts for worker quality. Another reason pertains to the only modest increase in value added relative to the increase in the wage bill and the number of employees decreased considerably over the sample period. From this it follows that over time the ratio of value added to the wage bill decreased, and that labour productivity and the cost per worker increased. If these increasing costs per worker reflected an increasing quality of the workforce we ought to use the wage bill as it adjusts for quality differences. Looking into our data, however, we do not find this to be the case as there are no considerable changes in the composition of the workforce.

free of adjustment costs.^{**} The Olley and Pakes (1996) (henceforth, OP) and labour productivity proxy estimators seek to consistently estimate the production function parameters by controlling (or proxying) for productivity using an investment policy function in the case of the OP decomposition, and a static intermediate input demand equation in the case of labour productivity. While both Olley and Pakes (1996) and labour productivity estimate the parameter on the variable input to production in the first stage, conditional on productivity, the ACF estimator dispenses with estimating the parameter on labour in the first stage because it is unlikely to be identified.^{††} The first stage is solely run to obtain an estimate of expected output $\widehat{\phi}_{it}$, which is then used in the second stage to identify the parameters on the inputs.

$$y_{it} = \underbrace{\alpha_l l_{it} + \alpha_k k_{it} + d^{-1}(d_{it}, k_{it})}_{\phi_{it}(l_{it}, k_{it}, d_{it})} + \epsilon_{it}$$

$$y_{it} = \alpha_l l_{it} + \alpha_k k_{it} + g(\phi_{it-1} - \alpha_l l_{it-1} - \alpha_k k_{it-1}) + \xi_{it} + \epsilon_{it}$$

where d^{-1} is the proxy function that is used to control for productivity. We follow the approach of Levinsohn and Petrin (2003) and use material inputs instead of investment in the proxy function. ξ_{it} is the innovation in the law of motion for productivity $\varphi_{it} = g_t(\varphi_{it-1}) + \xi_{it}$.

Consequently, productivity is assumed to follow an autoregressive Markov process, implying that, conditional on the exogenous productivity innovation, future productivity is affected only by past productivity. From the first stage we have an estimate of predicted output $\widehat{\phi}_{it}$ and can compute α_l and α_k using $\varphi_{it}(\alpha) = \widehat{\phi}_{it} - \alpha_l l_{it} - \alpha_k k_{it}$.

Relying on the Markov process of productivity we can recover $\xi_{it}(\alpha)$ as the residual by non-parametrically regressing $\varphi_{it}(\alpha)$ on its lag $\varphi_{it-1}(\alpha)$.^{‡‡} and form the following moments to obtain our estimates of the production function:

$$E \left\{ \xi_{it}(\alpha) \begin{pmatrix} l_{it} \\ k_{it} \end{pmatrix} \right\} = 0$$

Because we assume that both capital and labour are dynamic inputs to production, the innovation $\xi_{it}(\alpha)$ is independent of current capital, k_{it} , and current labour, l_{it} .

^{**} Hence, typically, capital can only be adjusted by the firm one period after it has experienced a shock, whereas labour can be adjusted instantaneously.

^{††} In both cases, it is critical that these proxy functions are monotonically increasing in φ_{it} so that they can be inverted and proxy for productivity.

^{‡‡} In practice, we rely on a third-order polynomial in lagged productivity to model future productivity flexibly. It is at this stage that we see that the assumption that productivity follows an autoregressive Markov process is crucial for identification.