Routledge
Taylor & Francis Group

Check for updates

On the Short- and Medium-Term Effects of Formalisation: Panel Evidence from Vietnam

AMADOU BOLY

Macroeconomics Policy, Forecasting and Research Department, African Development Bank (AfDB), Abidjan, Cote d'Ivoire

ABSTRACT *This paper analyses the consequences of formalisation on the performance of informal firms, using a panel dataset from Vietnam. We find that switching firms (before switching) have higher profit and value added compared to non-switching firms; suggesting heterogeneity. Becoming formal leads to an additional increase in switching firms' profit and value added. The benefits of formalisation materialise in the short-term (one year) and persist in the longer-term (three or more years). These benefits run through various channels such as better access to powered equipment or higher business association membership; but not better access to credit.*

Introduction

Three main views of informality can be found in the literature: dual, legalist, and exit views.[1] For the dual (or exclusion) view, informal firms are part of a subsistence sector and would be incapable of surviving in the formal sector, which they do not threaten (for example Harris & Todaro, 1970; La Porta & Shleifer, 2014). As a country develops, the efficient formal sector eventually crowds out the inefficient informal sector. The legalist view sees informality as the result of burdensome regulations that prevent small high-potential firms from entering the formal sector (for example de Soto, 1989).[2] The exit view of informality, which can be seen as more encompassing compared to the legalist one, suggests that informality stems from a deliberate private decision, after cost–benefit analysis by firms (see for example de Mel, McKenzie, & Woodruff, 2011; Maloney, 2004).[3] In the latter case, efforts to uncover the positive effects of formalisation can create incentives for firms to shift out of informality (Rand & Torm, 2012) and provide supporting evidence to policy-makers for promoting formalisation.

Several reasons may explain a firm's decision to formalise or not, including ignorance of legal obligations, initial reason for setting up a business (whether it is a real choice or by constraint), protection from corruption, expected access to market and larger business orders, or access to better location (see Cling, Razafindrakoto, & Roubaud, 2012 on the Vietnamese case).[4] In Hanoi and Ho Chi Minh City (Vietnam), over 80 per cent of formal firms see registration as beneficial, while nearly 50 per cent of informal firms see no value to it (Cling et al., 2012). Such a large difference between the real advantages of registration (given by formal firms) and the perceived benefits of registration for informal firms provides additional ground for trying to unveil the potential benefits of formalisation.

However, analysing the effects of formalisation on the performance of existing informal firms has been challenging due to unobserved heterogeneity. Indeed, firms choosing to formalise may have different underlying characteristics, such as the owner's abilities, business practices or firm preferences, compared to the ones that remained informal. A first objective of this paper is to analyse the consequences of formalisation on the performance of informal firms opting out of informality, while accounting for unobserved heterogeneity (which may explain self-selection into the formal sector). In this regard, it is worth stressing that this paper is not analysing the causes of formalisation.[5] Second, we analyse the effects of formalisation over time to show that they are long-lasting; an aspect that is absent in most previous studies. Finally, we look at some channels (discussed in the literature) that could help explain the impact of formalisation on firm performance; looking at the situation before and after switching.

This paper uses a panel dataset constructed from five small and medium enterprises (SME) surveys in Vietnam, conducted over the period 2005–2013. The dataset remains unique by the number of survey years (five), the number of firms, and its focus on the informal sector. We define formal firms as those that are registered to pay taxes (that is, have a tax code), a common indicator of formality in the literature (Fajnzylber, Maloney, & Montes Rojas, 2009; McKenzie & Sakho, 2010; Rand & Torm, 2012). Via the formal status variable (*Status* equals 0 if a firm is informal, and 1 if formal), we construct a variable, *Switcher*, which equals 1 for all years in which a firm that opted out of the informal sector has been observed in our panel dataset, irrespective of the year it became formal; and 0 if the firm remained informal throughout the survey periods. This dummy variable allows us to differentiate between always informal and switching firms explicitly.

We find that switching firms have higher profit and value added compared to informal non-switching firms. Such heterogeneity is typically assumed but not assessed in most previous studies. Becoming formal leads to a further increase in switching firms' profit and value added. Specifically, formalisation increases switchers' profit and value added respectively by 11.0 per cent and 8.9 per cent, compared to when they were informal. The benefits of formalisation exist in the short-term (one year) and persist over the longer-term (three or more years). These benefits run through channels such as better access to powered equipment, increased customer base, more advertising, and higher business association membership. However, we find no evidence of increased access to credit or increased likelihood to apply for formal loans.

The remainder of this paper is organised as follows. Section 2 briefly presents an overview of the existing literature on the impact of formalisation. In Section 3, we describe the dataset. Section 4 discusses the econometric approach, while Section 5 presents the main empirical results. We conclude in Section 6.

Literature review

The literature on the consequences of formalisation on firm performance mainly uses cross-sectional data and have relied on one or a combination of methods such as difference-in-differences, matching, instrumental variables, or regression discontinuity. The majority of these studies find that formalisation has a positive impact on firm performance (see for example Fajnzylber, Maloney, & Montes Rojas, 2011; McKenzie & Sakho, 2010; Rand & Torm, 2012).

Using firm-level cross-section data from Mexico, Fajnzylber et al. (2009) show that being formal increases profit by at least 20 per cent. Their approaches rely on matching under the assumptions that formal status is determined by a set of observable variables and on a control function approach. However, if selection into formality is based partly on unobserved characteristics, this may lead to overestimating the effects of formalisation (McKenzie & Sakho, 2010). Fajnzylber et al. (2011) used regression discontinuity and difference-in-differences to compare firms that were created immediately before and after a business tax reduction and simplification scheme (SIMPLES) in Brazil. They found that this reform led to increased levels of registration and to higher revenue, profit and employment among registered firms. As Fajnzylber et al. (2011) concentrate on newly created firms that opt for operating formally, not existing informal sector firms, the results can simply reflect self-selection at formal sector entry. McKenzie and Sakho (2010) estimate the impact of tax registration on firm profit in Bolivia, by using the distance between firm and tax office as an

instrument for registration status. The assumption is that being closer to a tax office increases the probability of registration. They find that the overall impact of tax registration is positive but heterogeneous; it leads to higher profits for medium-size firms in their sample, but has a negative impact on small and large firms. They also find that owners of larger informal firms have higher entrepreneurial abilities than owners of larger formal firms, in contrast to the mainstream view (see for instance La Porta & Shleifer, 2008). Using data from the 2006 World Bank survey of Indian microenterprises, Sharma (2014) finds, through propensity score matching, that registration leads to significant gains in sales per employee and value added per employee in India. An exception to the use of cross-section data is Rand and Torm (2012) who use a matched double-difference with the same panel data as in this study, but for 2007 and 2009 only. They find that registration leads to an increase in firm profits, investments and access to credit for Vietnamese SMEs; and to a decrease in the use of casual labour, indicating higher compliance with labour regulations. Compared to Rand and Torm (2012), the present study extends the panel dataset up to five observations per firm. As a result, we are able to estimate the effects of formalisation in the short but also in the medium-term; an aspect that is missing in Rand and Torm's study. Finally, De Mel, McKenzie, and Woodruff (2013) provide experimental evidence suggesting that firms become formal as the related benefits increase. In a field experiment in Sri Lanka, these authors randomly assigned 520 firms to five groups. In follow-up surveys, firms that formalised were found to have higher profits, but this result was driven by a few fast-growing firms: formalising had no effect on the profits of the majority.

Relative to transmission channels, formalisation is assumed to benefit the firms through increased access to credit, greater opportunities to engage with large firms and the government, or greater access to training and support programmes (Joshi, Prichard, & Heady, 2012); but the existing evidence is mixed. McKenzie and Sakho (2010) find that higher profits due to registration appear to come mainly from increase in customer base; and there was no impact of formalisation on the prospect of obtaining a bank loan. Fajnzylber et al. (2011) argue that improvements occur, not through access to credit or contracts with larger firms, but through lower cost of contracting labour, leading to the adoption of production techniques involving a permanent location and a larger paid labour force. Likewise, Rand and Torm (2012) could not obtain decisive evidence on the positive impact of formalisation on access to credit. De Andrade, Bruhn, and McKenzie (2016) find that registration increases advertising and use of receipt books, but not the likelihood of receiving government contracts, of using bank accounts or loans, or of participating in government programmes. According to Bruhn and McKenzie (2014), the likely explanation is that many informal firms would not receive credit or are unlikely to sell to the government anyway, even if they did register; and those firms that are in a position to do so, formalise when this need arises. Regarding the transmission channels, a distinctive aspect of this paper, relative to previous studies, is to compare the situation before and after formalisation, in order to ascertain that formalisation really had an impact on the transmission channel under consideration.

Data

Our dataset comes from SME surveys conducted in Vietnam in 2005, 2007, 2009, 2011, and 2013. The surveys, conducted by the Central Institute for Economic Management and the University of Copenhagen, cover about 2500 firms in each year. They were carried out in 10 locations; namely the cities of Hanoi, Hai Phong and Ho Chi Minh City, and rural provinces of Ha Tay, Phu Tho, Nghe An, Quang Nam, Khanh Hoa, Lam Dong, and Long An.

The population of non-state manufacturing enterprises was based on two data sources from the General Statistics Office of Vietnam (GSO): the Establishment Census from 2002 (GSO 2004) and the Industrial Survey 2004–2006 (GSO 2008). A representative sample of registered household and non-household firms was drawn from this population, using a stratified sampling procedure. The aim was to ensure the inclusion of an adequate number of enterprises in each province with different ownership forms, such as officially registered households, private firms, co-operatives, or limited liability companies. For reasons of implementation, the survey was confined to specific areas in each

province/city. In addition, the GSO enterprise census focused only on 'visible' firms (those with fixed professional premises), which resulted in an underestimation of household firms.

Informal household firms were included in the SME survey based on random on-site identification within the survey districts observed by the enumerator. With such an identification approach, the informal firms included in the survey are those operating alongside officially registered enterprises. These informal firms may be relatively more competitive (and profitable) compared to informal firms clustering in areas with none or very few formal firms (see Rand & Torm, 2012). In this regard, the sample of informal firms may not be fully representative of the informal sector as a whole in Vietnam.

Our total sample of firms with at least two observations includes always formal firms (1989 or 59.5% of total), switchers (458 or 13.7% of total) and always informal firms (896 or 26.8% of total).[6] For this study, we restrict the sample to informal or switching firms with at least two observations, resulting in a total of about 4800 observations (1300 firms) in the dataset. Out of these 1300 firms, 551 firms (42%) have observations in all five surveys, 173 firms (13%) have four observations, 294 (23%) have three observations, and 282 (22%) have two observations. In a subsequent section, we check the robustness of our results to attrition bias given its significance. As shown in Table 1, the sample is dominated by informal non-switching firms, which account for 67 per cent of the total number of firms; with switchers thus accounting for 33 per cent.

Empirical approach

To examine the effects of formalisation on firm performance, we exploit the panel nature of our dataset to estimate a random-effects (RE) model:

$$ln(y_{it}) = \rho_S D_i^S + \alpha F_{it} + \beta X_{it} + \lambda_t + \mu_i + \varepsilon_{it} \tag{1}$$

The dependent variable y_{it} represents two performance indicators, namely total gross profit and total value added. Value added can be defined as the value of a firm's output minus the costs of inputs used; such as raw materials, utilities (for example water, electricity or fuels) or services (repair and maintenance). Value added conventionally consists of labour income and profit. It thus constitutes a broader measure of firms' economic performance than profit which measures only returns to firm or capital owners (although these two variables can be the same, for example for a self-employed entrepreneur). Both value added and profit are used in this paper for robustness purposes. F_{it} is a dummy variable that indicates whether or not a firm is formal (0 if informal).

As highlighted in the introduction, one key difficulty in identifying the impact of formalisation on firm performance is that switching and always informal firms may not be comparable due to unobserved firm heterogeneity that is most likely to determine which firms switch and which firms do not. To address this issue, we control for firm-type fixed effects by using the variable F_{it} (0 if a firm is informal, and 1 if the firm is formal) to construct a variable, *Switcher*, denoted by D_i^S. If a firm has shifted out of the informal sector, the *Switcher* variable equals 1 for all years in which the firm has been observed in our panel dataset, irrespective of the year the firm switched; 0 if the firm remained informal throughout the five surveys. The inclusion of firm-type fixed effects in our regression model

Table 1. Frequency of firm types

Firm type	Overall		Between	
	Freq.	%	Freq.	%
Informal non-switcher	3131	64	869	67
Switcher (informal to formal)	1762	36	431	33
Total	**4893**	**100**	**1300**	**100**

(using dummy variable D_i^S) enables us to account explicitly for time-invariant unobserved heterogeneity between always informal and switching firms. Specifically, D_i^S captures unobserved differences between the switchers and the informal non-switchers (baseline group), while the variable F_{it} (0 if a firm is informal, and 1 if the firm is formal) captures the net effect of formalisation on firm performance.[7]

It can be noted that the use of random effects (instead of fixed effects model) in our regression analysis is driven by the fact that the variable *Switcher* is time-constant (that is being an always informal or a switching firm). A possible downside of random effects modelling relates to the requirement that the firm-specific effect (μ_i) be uncorrelated with the explanatory variables. As a result, for robustness check, we also use Mundlak's approach to correct for possible violation of the independence assumption between the covariates and the error term in the random effects model; through the inclusion of panel-group means of time-varying variables (see for example Bell & Jones, 2015; Mundlak, 1978).[8] We additionally estimate a fixed-effects (FE) model using the following equation for robustness check:

$$ln(y_{it}) = \alpha F_{it} + \beta X_{it} + \lambda_t + \mu_i + \varepsilon_{it} \tag{2}$$

The FE model controls for all time-invariant differences between the firms, so that the estimated coefficients cannot be biased because of omitted time-invariant characteristics.

In addition to the variables described above, we consider several control variables derived mainly from Rand and Torm (2012). The control variables include the gender of the owner/manager (0 if male, 1 otherwise); the education level of the owner/manager (0 if secondary school not completed, 1 otherwise) as a proxy for owner/manager's human capital; the number of regular full-time employees (in log), as well as the square, to control for firm size effects (McKenzie & Sakho, 2010); the share of production and service workers (as opposed to white-collar workers) to control for the average skill level in the firm, which can have an impact on firm performance (Rand & Torm, 2012); the share of female workers, which has been shown to depress wage levels in firms, thereby affecting performance (Larsen, Rand, & Torm, 2011); whether or not a firm owns a Certificate of Land Use Right (CLUR) to proxy property rights[9]; government inspection visits (0 if the firm has received no inspection in a given year, 1 if the number of inspection is equal to or more than 1); and dummy variables to control for industry, location, and time factors. The industry dummy variable equals 0 if the firm is in low-technology manufacturing, and 1 if the firm is in the medium-low or medium-high technology category.[10] Location dummies account for the fact that Vietnamese provinces are relatively autonomous, and have implemented centrally planned initiatives with varying degrees of speed and enthusiasm (Nguyen, Albrecht, Vroman, & Westbrook, 2007; Rand & Torm, 2012). Time dummies are included to control for potential time effects.

Results

In this section, we present results relative to the impact of formalisation on firm performance, as well as the time effects of formalisation. Note that the below discussion is based on estimates from the RE regression. We end this section by discussing possible transmission channels.

Dependent and control variables

Table 2 describes the dependent and independent variables per firm type (always informal or switching). We find that the average profit and value added of switchers is significantly higher than that of informal non-switchers (at the 1% level). Switchers are also significantly more likely to have a higher share of production workers; employ a higher number of full-time workers; receive more government inspections; and have an owner/manager who has completed secondary school. In contrast, informal

Table 2. Summary statistics of dependent and independent variables, by firm types

Variables	Informal non switcher			Switcher (informal to formal)		
	N	Mean	Sd	N	Mean	Sd
Dependent						
Profit, (log, real 1000 VND)	869	9.5	0.77	431	10.26	0.76
Value added, (log, real 1000 VND)	869	9.75	0.95	431	10.78	0.88
Independent						
Share of female employees	869	0.41	0.26	431	0.33	0.25
Share of production workers	869	0.61	0.19	431	0.7	0.15
Firm size (number of full time workers)	869	3.37	2.96	431	6.66	6.07
Gender of owner/manager (female = 0, male = 1)	869	0.32	0.39	431	0.27	0.35
Own land use right certificate (no = 0, yes = 1)	869	0.73	0.35	431	0.69	0.37
Compliance inspections (no = 0, yes = 1)	869	0.13	0.2	431	0.31	0.27
Owner/manager completed secondary school	869	0.86	0.26	431	0.89	0.22
Medium high-tech sector dummy	869	0.22	0.4	431	0.33	0.46

Note: The time-series average of each variable is first calculated by firm, before the average group statistics are computed.

Table 3. Summary statistics of dependent and independent variables, switchers only

Variables	Switchers before formalisation			Switchers after formalisation			Mean Difference
	N	Mean	Sd	N	Mean	Sd	
Dependent							
Profit, (log, real 1000 VND)	431	10.10	0.86	431	10.36	0.86	0.26***
Value added, (log, real 1000 VND)	431	10.63	0.95	431	10.87	0.97	0.24***
Independent							
Share of female employees	431	0.33	0.27	431	0.33	0.26	0.01
Share of production workers	431	0.77	0.18	431	0.66	0.18	−0.11***
Firm size (number of full time workers)	431	7.08	7.57	431	6.19	5.62	−0.89*
Gender of owner/manager (female = 0, male = 1)	431	0.25	0.41	431	0.28	0.40	0.03
Own land use right certificate (no = 0, yes = 1)	422	0.67	0.44	431	0.70	0.41	0.03
Compliance inspections (no = 0, yes = 1)	431	0.38	0.43	431	0.24	0.34	−0.14***
Owner/manager completed secondary school	431	0.88	0.30	431	0.89	0.26	0.02
Medium high-tech sector dummy	431	0.33	0.47	431	0.32	0.46	0.00

Note: The time-series average of each variable is first calculated by firm, before the average group statistics are computed. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

non-switchers are significantly more likely than switchers to have a higher share of female employees, a female owner/manager, and a CLUR.

Table 3 shows the same variables as in Table 2, but only for switchers, before and after switching. It shows that the average profit and value added of switchers increase after formalisation. Relative to control variables, we find that formalisation leads to a significant reduction in the share of production workers and the firm size. Such a result suggests that formalisation may lead to downsizing, in order to increase labour productivity in the switching firm. Formalisation also leads to a significant reduction in the likelihood of compliance inspections; likely because formalised firms are no longer operating in informality. Nevertheless, these formalised firms remain more likely to receive compliance visits compared to always-informal firms, as they are on average larger and thus more exposed.

Table 4. Effects of formality on profit and value added

Variables	Log profit (real 1,000 VND)			Log Value Added (real 1,000 VND)		
	RE	Mundlak	FE	RE	Mundlak	FE
Switcher (from informal to formal)	0.20***	0.24***		0.28***	0.28***	
	(0.04)	(0.08)		(0.04)	(0.07)	
Switcher (after formalisation)	0.11**	0.12***	0.14***	0.09***	0.10***	0.11***
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Share of female employees	−0.16***	0.05	0.05	−0.28***	−0.03	−0.02
	(0.05)	(0.07)	(0.08)	(0.05)	(0.06)	(0.07)
Share of production workers	0.07	0.16***	0.14**	0.15***	0.20***	0.19***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)
Firm size (log (1+employement))	1.39***	1.02***	1.01***	1.79***	1.38***	1.38***
	(0.09)	(0.10)	(0.12)	(0.09)	(0.09)	(0.11)
Firm size square (log (1+employment))	−0.15***	−0.10***	−0.10***	−0.16***	−0.12***	−0.12***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
Gender of owner/manager (male = 0, female = 1)	−0.04	−0.01	−0.02	−0.02	0.02	−0.01
	(0.03)	(0.03)	(0.04)	(0.03)	(0.02)	(0.03)
Own CLUR (no = 0, yes = 1)	0.05*	0.05**	0.04	0.00	0.01	0.01
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
Compliance inspections (no = 0, yes = 1)	0.17***	0.16***	0.15***	0.15***	0.13***	0.12***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Owner/manager completed secondary school	0.07**	0.06*	−0.02	0.09***	0.08***	−0.00
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)
Medium-high tech sector dummy	−0.04	−0.09***	−0.02	0.08***	−0.00	0.06
	(0.03)	(0.03)	(0.11)	(0.03)	(0.03)	(0.09)
Constant	8.06***	7.58***	8.08***	7.76***	7.15***	7.91***
	(0.14)	(0.14)	(0.13)	(0.13)	(0.13)	(0.12)
Observations	**4,792**	**4,792**	**4,792**	**4,792**	**4,792**	**4,792**
R-squared			**0.16**			**0.30**
Time dummies included	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Province dummies included	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Number of panels	**1,300**	**1,300**	**1,300**	**1,300**	**1,300**	**1,300**

Notes: Robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. In Mundlak's model, the panel mean of independent variables is included in the regression, except: Switcher (from informal to formal).

Impact of formalisation on profit and value added

A first objective of this paper is to analyse the effects of formalisation on the performance of informal firms opting out of informality. In this regard, Table 4 shows RE, Mundlak and FE regressions for (log) total gross profit and (log) value added. Note that the results on the effects of formalisation do not change qualitatively if controls are introduced progressively.[11]

Our first result provides evidence that switching firms are different from informal non-switching firms. Looking at the coefficient of *Switcher* (RE columns), we find that the profit and value added levels of switchers are significantly higher respectively by 21.8 per cent and 31.6 per cent compared to those of informal non-switchers.[12] In most of the previous studies, such a difference was assumed but not assessed. Thus, even before switching to the formal sector, switchers have higher profit and value added compared to informal non-switchers. Ignoring this difference (for example in OLS regressions) would lead to an overestimation of the effects of formalisation for switching firms.

Our second result indicates that becoming formal leads to an increase in profit and value added, as shown by the coefficient of 'Switcher (after formalisation)'; and this can be considered the 'net' effect of formalisation on switchers. Formalisation increases total amount of profit and

value added of switchers significantly, by 11.0 per cent and 8.9 per cent respectively, compared to when they were informal.[13] The results are comparable with either the RE, Mundlak or FE estimations.

Several other control variables are noteworthy in Table 4. First, firm size has a significant positive impact on the amount of profit and value added; but at a decreasing rate. Second, the gender of the owner/manager is typically not a significant determinant of firm performance. Third, receiving at least one compliance inspection is positively related to profit and value added. This is possible if inspections lead to increased firms' compliance with labour regulations, forcing them to reduce size (as found in Table 3). Compliance with labour regulations and size reduction result in keeping only the most productive workers, thereby increasing productivity and profits (see also Rand & Torm, 2012). Whether the owner or manager of the firm has completed secondary school matters positively; highlighting the importance of human capital and skills.

Robustness check to endogeneity

In this section, we analyse the potential endogeneity of formalisation using a Control Function approach (see Wooldridge, 2015 for details). The Control Function approach first estimate a model of the endogenous explanatory variable (see Equation [3] below), then uses the estimated model to obtain 'generalised residuals' (see Equation [4] below) that are added as an additional regressor in the main regression (that is Equation [1]). The approach relies on the usual requirement that there be at least one exogenous variable that is omitted from Equation (1) and that is partially correlated with the dependent variable in Equation (3) (see Wooldridge, 2015).

In a first step, we therefore estimate a clustered Probit model based on the following equation:

$$F_{it} = 1[\beta X_{it} + \delta I_{it} + v_{it}] = 1[\omega Z_{it} + v_{it}] \tag{3}$$

Where $1[.]$ is the binary indicator function; F_{it}, the dependent variable, is a dummy variable which takes value 1 if a firm is formal, and 0 otherwise; X_{it} are control variables described earlier in Section 4; I_{it} corresponds to a set of exogenous variables that are omitted from Equation (1), and that are partially correlated with formalisation; $Z_{it} = (X_{it}, I_{it})$; and v_{it} is an error term.

To construct I_{it}, we compute the annual provincial-level averages for each of the three following binary variables: access to powered equipment (1 if access, 0 otherwise); compliance visits (0 if the firm has received no inspection in a given year, 1 if the number of inspection is equal to or more than one); and bribe payments (1 if the firm has made any bribe payments in a given year, 0 otherwise). We use only always-formal firms and formalised firms (that is switchers); except that the latter are excluded from calculations for the year they became formal. As mentioned in Section 3, informal firms in the sample are selected based on random on-site identification and are those operating alongside formal firms. Our assumption is that these informal firms are more likely to formalise when they are able to observe some characteristics of formal firms and presumably attribute those characteristics to formalisation.

The results are shown in column 'Clustered Probit' of Section 1 in the Supplementary Materials. They suggest that compliance visits received and bribes paid by formal firms have negative and significant effects on the likelihood of formalisation; while access to powered equipment has a positive and significant impact. As a result, I_{it} meets the usual requirement that there be at least one exogenous variable that is omitted from the main equation and that is partially correlated with the dependent variable in Equation (3).

In a second step, the results from Equation (3) are used to obtain 'generalised residuals' as:

$$\hat{r}_{it} = F_{it}\lambda\left(\hat{\delta}Z_{it}\right) - (1 - F_{it})\lambda\left(-\hat{\delta}Z_{it}\right) \tag{4}$$

Finally, we re-estimated Equation (1) while adding $\hat{r}_{it}$ as a regressor to control for endogeneity. The results are presented in column 'RE model' of Section 1 of the Supplementary Materials and confirm that formalisation has a positive impact on switchers' profit and value added, compared to informal firms.

Robustness check to attrition bias

As mentioned in Section 3, there is significant attrition in our panel data. As a result, we checked the robustness of our results to attrition bias, a typical challenge in panel datasets.[14] We do so first by varying the sample size, then using a test suggested by Verbeek and Nijman ([1992]) and another one suggested by Becketti, Gould, Lillard, and Welch ([1988]).[15] Together, these checks suggest that our results are robust to attrition bias.

As a first check, we vary the sample size by using the full sample or keeping only firms with at least three, four, or five observations; noting that the results discussed in this paper are based on firms with at least two observations. The full model is estimated for each sample size but we report the coefficients in Table 5 for only the variables 'Switcher (from informal to formal)' and 'Switcher (after formalisation)' to save space. Whether the sample includes all firms, the balanced panel or only firms with at least three or four observations, we find that switching firms are different from informal non-switchers and the impact of formalisation is positive and significant at conventional levels.

The test suggested by Verbeek and Nijman ([1992]) consists of adding an attrition variable (Attrition = 0 if a firm is observed in all five survey waves; 1 otherwise) to the explanatory variables. The coefficient of this variable (Attrition) is insignificant in all regressions, suggesting that attrition is ignorable as shown in Section 2 of the Supplementary Materials. In order to carry out a statistical test of the difference in coefficients between attrition firms and remaining firms, we follow Becketti, Gould, Lillard and Welch's ([1988]) approach by creating interaction variables between the attrition variable (Attrition) and all other explanatory variables, except year dummies. Equation (1) is then re-estimated, adding the 'Attrition' variable plus its interactions with the other explanatory variables as regressors (see Section 3 in the Supplementary Materials). First, it can be noted that the Attrition variable is insignificant for both profit (p-value = 0.63) and value added (p-value = 0.15). Second, the interaction variable 'Attrition × Switcher (from informal to formal)' is found not to be significantly different from 0. The p-values are 0.30 and 0.24 for profit and value added respectively. These results suggest that switchers that leave our panel sample are not significantly different from those that do not. Likewise, the variable 'Attrition × Switcher (after formalisation)' is not significantly different from 0. For profit, the p-value is 0.25; while for value added, the p-value is 0.35. The effects of formalisation on performance are thus not significantly different between attrition firms and remaining firms.

Persistence of formalisation effects

A second objective of this study is to analyse the effects of formalisation over time, an aspect that is absent in most previous studies. Given that registration costs can affect performance (negatively) in the period immediately following formalisation, the potential benefits of formalisation can materialise with a delay. It is therefore important to analyse both the short-term and longer-term effects of formalisation. Such an analysis can be valuable for firms that are considering to formalise as well as policymakers trying to promote formalisation.

To analyse persistence, we constructed dummies for specific lengths of time (one, three, and five or more) which measures the number of years since a firm has shifted out of the informal sector. In the first survey year (2005), all firms were informal. The first switchers are recorded in 2007, with formalisation having taken place between 2005 and 2007. For all switching firms, we assume that formalisation took place in the year between two surveys. As a result, for firms that were informal in 2005 but formal in 2007, the year of formalisation is set at 2006; and the numbers of years since switching is one in 2007. For these firms that switched in 2007, the number of years since switching becomes three in 2009, five in 2011, and seven in 2013.[16] For firms that switched in 2009, the number of years since switching is one in 2009, three in 2011, and five in 2013. Finally, for firms that became

Table 5. Effects of formality on profit and value added – varying sample sizes

Variables	Log Profit (real 1000 VND)			Log Value Added (real 1000 VND)		
	RE	Mundlak	FE	RE	Mundlak	FE
Full Sample						
Switcher (from informal to formal)	0.18***	0.24***		0.25***	0.28***	
	(0.04)	(0.08)		(0.03)	(0.07)	
Switcher (after formalisation)	0.10**	0.12***	0.14***	0.08***	0.09***	0.11***
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Observations	**5,159**	**5,159**	**5,159**	**5,159**	**5,159**	**5,159**
Number of panels	**1,667**	**1,667**	**1,667**	**1,667**	**1,667**	**1,667**
At least 3 observations						
Switcher (from informal to formal)	0.19***	0.25***		0.30***	0.28***	
	(0.05)	(0.08)		(0.04)	(0.08)	
Switcher (after formalisation)	0.11**	0.13***	0.14***	0.08**	0.09***	0.10***
	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
Observations	**4,239**	**4,239**	**4,239**	**4,239**	**4,239**	**4,239**
Number of panels	**1,018**	**1,018**	**1,018**	**1,018**	**1,018**	**1,018**
At least 4 observations						
Switcher (from informal to formal)	0.21***	0.30***		0.32***	0.34***	
	(0.05)	(0.09)		(0.04)	(0.08)	
Switcher (after formalisation)	0.11**	0.14***	0.14***	0.07*	0.10***	0.10**
	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)
Observations	**3,375**	**3,375**	**3,375**	**3,375**	**3,375**	**3,375**
Number of panels	**724**	**724**	**724**	**724**	**724**	**724**
Balanced panel (5 observations)						
Switcher (from informal to formal)	0.24***	0.27***		0.32***	0.25***	
	(0.05)	(0.10)		(0.05)	(0.09)	
Switcher (after formalisation)	0.09*	0.11**	0.11**	0.07*	0.09**	0.09**
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)
Observations	**2,704**	**2,704**	**2,704**	**2,704**	**2,704**	**2,704**
Number of panels	**551**	**551**	**551**	**551**	**551**	**551**
Controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Time dummies included	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Province dummies included	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ The full model is estimated for each sample size but the coefficients of the following variables are not reported to save space: Share of female employees; Share of production workers; Firm size (log (1+employement)); Firm size square (log (1+employment)); Gender of owner/manager (male = 0, female = 1); Own land use right certificate, CLUR (no = 0, yes = 1); Compliance inspections (no = 0, yes = 1); Owner/manager completed secondary school; Medium high-tech sector dummy.

formal in 2011, the number of years since switching is one in 2011 and three in 2013, the year of the last survey. In total, the number of firms that switched from the informal to the formal sector for one, three, and five years or more is, respectively, 431, 321, and 269.[17]

The results are shown in Table 6 for RE and FE models. For both profit and value added, we find that the coefficients are positive and significant for all lengths of time (one, three, and five years or more. Our results are therefore supportive of the fact that the benefits of formalisation materialise in the short-term (length one) and persist over time (length three and length five years or more). Using a control function approach to control for endogeneity of formalisation in the first year does not change the results (see Section 4 in the Supplementary Materials).

Table 6. Persistence of the effects of formality on profit and value added

Variables	RE		FE	
	Log Profit (real 1000 VND)	Log Value Added (real 1000 VND)	Log Profit (real 1000 VND)	Log Value Added (real 1000 VND)
Switcher (from informal to formal)	0.12*** (0.04)	0.18*** (0.04)		
Time since becoming formal (dummy, 1 years)	0.27*** (0.04)	0.25*** (0.03)	0.29*** (0.04)	0.27*** (0.03)
Time since becoming formal (dummy, 3years)	0.30*** (0.05)	0.31*** (0.04)	0.30*** (0.05)	0.31*** (0.04)
Time since becoming formal (dummy, 5 years)	0.32*** (0.05)	0.39*** (0.04)	0.35*** (0.06)	0.39*** (0.05)
Share of female employees	−0.18*** (0.05)	−0.29*** (0.05)	0.02 (0.08)	−0.05 (0.07)
Share of production workers	−0.15*** (0.05)	−0.11** (0.05)	−0.06 (0.05)	−0.06 (0.05)
Firm size (log (1+employement))	1.42*** (0.10)	1.84*** (0.09)	1.04*** (0.12)	1.40*** (0.12)
Firm size square (log (1+employment))	−0.16*** (0.02)	−0.17*** (0.02)	−0.10*** (0.03)	−0.12*** (0.03)
Gender of owner/manager (male = 0, female = 1)	−0.05 (0.03)	−0.03 (0.03)	−0.02 (0.04)	−0.01 (0.03)
Own land use right certificate, CLUR (no = 0, yes = 1)	0.05* (0.03)	0.01 (0.03)	0.05 (0.03)	0.02 (0.03)
Compliance inspections (no = 0, yes = 1)	0.08*** (0.03)	0.07*** (0.03)	0.08*** (0.03)	0.06** (0.03)
Owner/manager completed secondary school	0.08** (0.03)	0.10*** (0.03)	−0.02 (0.04)	−0.00 (0.04)
Medium high-tech sector dummy	−0.06* (0.03)	0.05* (0.03)	−0.03 (0.11)	0.04 (0.09)
Constant	8.39*** (0.13)	8.14*** (0.13)	8.33*** (0.12)	8.19*** (0.12)
Observations	**4,792**	**4,792**	**4,792**	**4,792**
R-squared			**0.13**	**0.26**
Time dummies included	**No**	**No**	**No**	**No**
Province dummies included	**Yes**	**Yes**	**Yes**	**Yes**
Number of panels	**1,300**	**1,300**	**1,300**	**1,300**

Note: Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

Transmission channels

There are a number of possible channels through which formality can have a positive impact on firm performance. These channels include access to credit, greater opportunities to engage with large firms and the government, greater access to training and support programmes (Joshi et al., 2012); the opportunity to enlarge customer base and lower the costs of corruption (McKenzie & Sakho, 2010); or the ability to lower the cost of contracting labour (Fajnzylber et al., 2011). Previous studies have typically considered these transmission channels as conduits to higher firm performance; but, they may also constitute reasons why a firm may have switched out of the informal sector.

By considering the periods before and after formalisation, our analysis is useful in understanding whether the benefits expected from formalisation materialise or not for switching firms. In this section, we analyse the following channels: access to better equipment, access to a larger customer pool, access to formal credit, increased likelihood of applying to formal loans, access to business association membership, more advertising or less bribe payments (see Table 7).

Switchers' probability of accessing powered equipment (or both manual and powered equipment) increases significantly when they shift out of the informal sector. This result is explicable if formalisation also leads to better access to public utilities such as electricity. Access to better equipment can can improve productivity, and thereby performance. Being able to attract more customers can boost sales and increase profits. Customer base is a dummy variable taking value 0 if less than 20; 1 otherwise. Switchers' customer base tends to be smaller than that of informal non-switching firms, before formalisation. After formalisation, the likelihood of having a larger customer base increases significantly. This suggests that some firms may switch out of the informal sector in order to alleviate the constraint of a small customer base. In contrast, Rand and Torm (2012) do not find any improvements along the customer base dimension of becoming formal in their study; while McKenzie and Sakho (2010) find a positive impact.

We find that the likelihood of accessing credit does not increase with formalisation; and this result is also found by Fajnzylber et al. (2011), McKenzie and Sakho (2010) and Rand and Torm (2012). Still, compared to previous studies, we find that switchers have a greater probability of accessing formal loans before formalising, compared to informal non-switchers. A similar result applies to formal loan applications: switchers are more likely to apply for formal loans than informal non-switchers, but only marginally (at the 10% level). Formalisation does not change their behaviour (as in de Andrade et al., 2016). It can be noted that in Vietnam, firms can use their CLUR to obtain formal credit, even when they are informal, partly explaining the previous results.

Our results suggest that switchers have an increased likelihood of being a member of a business association, which can introduce the entrepreneur to new technologies or ways of doing business (Fajnzylber et al., 2011), but only after they become formal (not before). The likelihood of advertising is higher for switchers, compared to informal non-switchers. This likelihood further increases after they shift out of the informal sector (see also de Andrade et al., 2016). Similarly, we find that switchers' probability of bribe payments is higher compared to informal firms and that formalisation leads to a further increase in the probability of paying bribes. This is in line with the fact that always formal firms (in our sample) have significantly higher bribe payments than informal firms but contrary to the assumption that firms may formalise to protect themselves from corruption (for example McKenzie & Sakho, 2010).

In summary, we find that formalisation can benefit informal firms through better access to improved equipment, larger customer base, advertising and business association membership. However, we find no evidence of increased likelihood of applying for formal loans or improved access to credit, a key reason often put forward to explain the formalisation of informal firms. Formalisation also appears to increase the probability of bribe payments.

Conclusion

Using a panel dataset consisting of five waves of SME surveys in Vietnam, this paper analyses the impact of formalisation on firm performance, the persistence of these effects in the longer-term, and the channels

Table 7. Transmission channels of the effects of formality (Logit)

Variables	Type of machinery	Size of customer base	Firm has access to formal credit	Firm has applied for formal credit	Firm is member of business association	Firm advertises	Firm pays bribe
Switcher (from informal to formal)	0.49* (0.29)	−0.53*** (0.12)	0.33** (0.16)	0.26* (0.16)	−0.28 (0.35)	0.77** (0.38)	0.25** (0.12)
Switcher (after formalisation)	0.57** (0.26)	0.59*** (0.12)	−0.11 (0.14)	−0.00 (0.14)	0.79*** (0.31)	1.02*** (0.29)	0.53*** (0.12)
Share of female employees	−2.03*** (0.33)	−0.08 (0.16)	−0.70*** (0.22)	−0.69*** (0.21)	0.50 (0.45)	0.96* (0.55)	−0.76*** (0.19)
Share of production workers	−0.61** (0.31)	−0.08 (0.17)	1.33*** (0.22)	1.25*** (0.22)	−1.46** (0.57)	−0.75 (0.70)	−0.21 (0.23)
Firm size (log (1+employment))	1.55*** (0.46)	−0.06 (0.27)	1.04*** (0.38)	1.06*** (0.36)	2.80*** (0.74)	0.96 (0.88)	2.04*** (0.35)
Firm size square (log (1+employment))	−0.29*** (0.13)	−0.01 (0.09)	−0.08 (0.12)	−0.10 (0.11)	−0.33** (0.15)	−0.03 (0.18)	−0.30*** (0.11)
Gender of owner/manager (male = 0, female = 1)	−0.21 (0.11)	−0.08 (0.07)	0.24** (0.09)	0.21* (0.09)	0.10 (0.15)	−0.53* (0.17)	−0.12 (0.08)
Own land use right certificate (no = 0, yes = 1)	0.03 (0.18)	0.18** (0.09)	−0.01 (0.12)	−0.05 (0.11)	0.01 (0.23)	0.36 (0.30)	0.04 (0.11)
Compliance inspections (no = 0, yes = 1)	0.62*** (0.17)	−0.60*** (0.09)	0.47*** (0.11)	0.46*** (0.10)	0.06 (0.23)	−0.34 (0.29)	0.21** (0.10)
Owner/manager completed secondary school	0.14 (0.18)	0.06 (0.10)	0.02 (0.11)	0.07 (0.11)	0.21 (0.25)	−0.28 (0.31)	0.30** (0.10)
Medium-high tech sector dummy	0.08 (0.20)	0.01 (0.11)	−0.01 (0.14)	0.01 (0.14)	−0.38 (0.33)	0.52* (0.32)	0.02 (0.15)
Constant	2.69*** (0.66)	−0.01 (0.33)	−4.32*** (0.47)	−4.06*** (0.45)	−8.37*** (0.91)	−7.55*** (1.29)	−3.25*** (0.37)
lnsig2u	2.10*** (0.13)	−0.31** (0.14)	0.68*** (0.11)	0.60*** (0.11)	1.14*** (0.21)	0.32 (0.36)	−1.70*** (0.49)
Observations	**4,760**	**4,703**	**4,792**	**4,784**	**4,518**	**4,792**	**4,792**
Time dummies included	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Province dummies included	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Number of panels	**1,295**	**1,299**	**1,300**	**1,300**	**1,226**	**1,300**	**1,300**

Note: Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

through which these impacts materialised. Such an analysis has been challenging because of potential selection bias, due to the fact that firms choosing to formalise can have different underlying characteristics, such as the owner's abilities or firm preferences, compared to those that remained informal.

To control for unobserved heterogeneity, we created a dummy variable that distinguishes between the two groups of firms: always informal and switchers. We find that switching firms perform better than informal non-switching firms. Such heterogeneity is typically assumed in most previous studies. Second, we find that becoming formal leads to a further increase in switching firms' profit and value added. Third, the benefits of formalisation materialise in the short-term and persist over time. Finally, we show that the benefits of formalisation run through channels such as better access to powered equipment, increased customer base, more advertising, and higher business association membership. Yet, formalisation does not seem to improve access to credit or increase the likelihood to apply for formal loans.

Our results are also broadly consistent with the hypothesis that firms rationally make the decision to formalise by comparing the costs and the benefits. The firms that formalise, on average, made the right decision. Consequently, the results of this study highlight the need for a policy mix that reduces the cost of registration, showcases the potential benefits of formalisation, and further increases the attractiveness of the formal sector. The latter strategy could put into place supportive measures to facilitate access to credit, to modern production equipment, or business associations; or reduce the incidence of corruption. Noteworthily, explicitly showing that there is a difference between informal firms that formalise and those that remain informal give, to some extent, support to the dual view which argues that some informal firms would be incapable of surviving in the formal sector. Those firms would not formalise whatever the incentives.

Acknowledgements

Disclosure statement

Notes

1. See for example Perry et al. (2007) or Cling et al. (2012).
2. One policy implication of the legalist view is the removal of costly entry regulations, but evidence suggests that ease of formalisation alone will not induce most informal firms to become formal (Bruhn & McKenzie, 2014).
3. For example, McKenzie and Sakho (2010) hypothesise that a profit-maximising firm becomes formal if and only if the expected present discounted value of the net benefits from doing so outweighs the upfront costs:

$$\sum_{t=1}^{T} \delta^t E\left(\pi_{F,t} - \pi_{I,t}\right) + \theta_{law-abiding} > C_{Money} + C_{Time} + C_{Information}$$

where $\pi_{F,t}$ denotes the firm's profits if it is formally registered at time t, and $\pi_{F,t}$ denotes the firm's profits if it is not formally registered at time t. $\theta_{law-abiding}$ denotes the utility benefit to firm owners from obeying the law and feeling they are contributing to national welfare through paying taxes. C_{Money}, C_{Time} and $C_{Information}$ denote the monetary, time, and information costs from registering.

4. Interestingly, no mention is made of past performance (such as past profit) as a driver for formalisation, suggesting that firms may be forward-looking when deciding to formalise.
5. See for example Cling et al. (2012) for details on the causes of formalisation in Vietnam.
6. We assume that once a firm becomes formal, it stays formal and recoded formality status accordingly. This applies to about 2.4 per cent of observations of the initial total sample. The main justification is that once a firm enters tax authorities' records by acquiring a tax code, it becomes very difficult for the firm to move again into informality.
7. An implicit assumption is that formalisation happens randomly within the Switchers group.
8. Bell and Jones (2015, p. 133) write the following: 'The downside of Random Effects (RE) modelling – correlated lower-level covariates and higher-level residuals – is omitted-variable bias, solvable with Mundlak's (1978) formulation. Consequently, RE can provide everything that FE promises and more'. Mundlak's approach is also discussed in some textbooks such as Baltagi (2008) or Wooldridge (2010).
9. Rand and Torm (2012) typically control for this variable in their empirical model, based on the fact that even informal firms in Vietnam are generally able to use their CLUR as collateral for a loan, thereby easing potential financial constraints for increased investments and performance.
10. We use the Organisation for Economic Co-operation and Development (OECD) technology classification.
11. Available upon request.
12. As we are using a semi-logarithmic functional form, we estimate the effect of a dummy variable coefficient on the dependent variable as: $g^* = \exp(\hat{c} - \frac{1}{2}V(\hat{c})) - 1$,
   where $\hat{c}$ is the dummy variable coefficient and $V(\hat{c})$ its variance (see Halvorsen & Palmquist, 1980; Kennedy, 1981). Because informal firms in the sample may be relatively more competitive (and profitable) compared to the informal sector population, these estimates can be seen as lower bounds.
13. See endnote 12.
14. One possibility could have been to limit the analysis to the balanced panel as done by Rand and Torm (2012).
15. See also Baulch and Quisumbing (2011).
16. These numbers can also be interpreted in terms of ranges, that is, between zero and two in 2007; between two and four in 2009; between four and six in 2011; and between six and eight in 2013.
17. We merge firms that have been formal for five or seven years into a single category: 'five years or more'.

References

Baltagi, B. (2008). *Econometric analysis of panel data*. Chichester, UK: John Wiley & Sons.
Baulch, B., & Quisumbing, A. (2011). *Testing and adjusting for attrition in household panel data* (Toolkit Note). Manchester: Chronic Poverty Research Centre.
Becketti, S., Gould, W., Lillard, L., & Welch, F. (1988). The panel study of income dynamics after fourteen years: An evaluation. *Journal of Labor Economics*, 6, 472–492. doi:10.1086/298192
Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, 3(1), 133–153. doi:10.1017/psrm.2014.7
Bruhn, M., & McKenzie, D. (2014). Entry regulation and the formalization of microenterprises in developing countries. *The World Bank Research Observer*, 29, 186–201. doi:10.1093/wbro/lku002
Cling, J.-P., Razafindrakoto, M., & Roubaud, F. (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, 632–652. doi:10.1080/13547860.2012.724553
De Andrade, G. H., Bruhn, M., & McKenzie, D. (2016). A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms. *World Bank Economic Review*, 30, 24–54. doi:10.1093/wber/lhu008
De Mel, S., McKenzie, D., & Woodruff, C. (2011). *What is the cost of formality? Experimentally estimating the demand for formalization* (Working Paper). Coventry: University of Warwick.
De Mel, S., McKenzie, D., & Woodruff, C. (2013). The demand for, and consequences of, formalization among informal firms in Sri Lanka. *American Economic Journal: Applied Economics*, 5(2), 122–150. doi:10.1257/app.5.2.122
De Soto, H. (1989). *The other path: The invisible revolution in the Third World*. New York, NY: Harper and Row.
Fajnzylber, P., Maloney, W., & Montes Rojas, G. (2009). Releasing constraints to growth or pushing on a string? Policies and performance of Mexican micro-firms. *The Journal of Development Studies*, 45, 1027–1047. doi:10.1080/00220380802264911
Fajnzylber, P., Maloney, W., & Montes Rojas, G. (2011). Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program. *Journal of Development Economics*, 94(2), 262–276. doi:10.1016/j.jdeveco.2010.01.009
General Statistics Office. (2004). *Results of establishment census of Vietnam 2002: Business establishments*. Hanoi: Statistical Publishing House.

General Statistics Office. (2008). *The real situation of enterprises through the results of surveys conducted in 2005, 2006, 2007*. Hanoi: Statistical Publishing House.

Halvorsen, R., & Palmquist, R. (1980). The interpretation of dummy variables in semi-logarithmic equations. *American Economic Review*, 70, 474–475.

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review*, 60, 126–142.

Joshi, A., Prichard, W., & Heady, C. (2012). *Taxing the informal economy: Challenges, possibilities and remaining questions* (ICTD Working Paper 4). Brighton: International Centre for Tax and Development.

Kennedy, P. E. (1981). Estimation with correctly interpreted dummy variables in semi-logarithmic equations. *American Economic Review*, 71, 801.

La Porta, R., & Shleifer, A. (2008). *The unofficial economy and economic development* (Brookings Papers on Economic Activity No. 2). Washington, DC: Brookings Institute.

La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, 28(3), 109–126. doi:10.1257/jep.28.3.109

Larsen, A. F., Rand, J., & Torm, N. (2011). Do recruitment ties affect wages? An analysis using matched employer-employee data from Vietnam. *Review of Development Economics*, 15, 541–555. doi:10.1111/j.1467-9361.2011.00626.x

Maloney, W. F. (2004). Informality revisited. *World Development*, 32, 1159–1178. doi:10.1016/j.worlddev.2004.01.008

McKenzie, D., & Sakho, Y. S. (2010). Does it pay firms to register for taxes? The impact of formality on firm profitability. *Journal of Development Economics*, 91(1), 15–24. doi:10.1016/j.jdeveco.2009.02.003

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, 46, 69–85. doi:10.2307/1913646

Nguyen, B. T., Albrecht, J. W., Vroman, S. B., & Westbrook, M. D. (2007). A quantile regression decomposition of urban-rural inequality in Vietnam. *Journal of Development Economics*, 83(2), 466–490. doi:10.1016/j.jdeveco.2006.04.006

Perry, G., Maloney, W., Arias, O., Fajnzylber, P., Mason, A., & Saavedra-Chanduvi, J. (2007). *Informality: Exit and exclusion*. Washington, DC: World Bank Publications.

Rand, J., & Torm, N. (2012). The benefits of formalization: Evidence from Vietnamese manufacturing SMEs. *World Development*, 40, 983–998. doi:10.1016/j.worlddev.2011.09.004

Sharma, S. (2014). Benefits of a registration policy for microenterprise performance in India. *Small Business Economics*, 42, 153–164. doi:10.1007/s11187-013-9475-y

Verbeek, M., & Nijman, T. (1992). Testing for selectivity bias in panel data models. *International Economic Review*, 33, 681–703. doi:10.2307/2527133

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

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, 50, 420–445. doi:10.3368/jhr.50.2.420