Empirical Economics (2022) 62:1773–1800
https://doi.org/10.1007/s00181-021-02072-w

Effects of tax administration corruption on innovation inputs and outputs: evidence from small and medium sized enterprises in Vietnam

Hung Quang Doan[1,2,3] · **Nam Hoang Vu**[2] · **Binh Tran-Nam**[4,5] · **Ngoc-Anh Nguyen**[3]

Received: 28 September 2020 / Accepted: 17 May 2021 / Published online: 5 June 2021

Abstract
This paper examines the influence of tax administration corruption on different types of innovation inputs and outputs in Vietnam. It utilizes firm-level panel data derived from biennial surveys of small and medium sized enterprises (SMEs) in Vietnam from 2005 to 2015. In terms of estimation method, the study applies the control function method for a dynamic binary response panel data model with endogenous explanatory variables, state dependence, and initial condition problems simultaneously. The key estimation results confirm the grease-the-wheels hypothesis that petty tax corruption positively affects all types of firm innovative activities. It is further found that innovation outputs and machinery innovation input of an SME are positively determined by its innovation 2 years earlier and innovation in the initial period. The key finding of the study implies that it is a challenge for governments in transition economies to fight against tax corruption, especially for Vietnam, which is known to be a high tax collection, high tax effort country.

Keywords Tax administration corruption · Innovation · Small and medium sized enterprises · Vietnam

✉ Binh Tran-Nam
b.tran-nam@unsw.edu.au

1 Department of Business and Management, LUISS University, Rome, Italy

2 Faculty of International Economics, Foreign Trade University, Hanoi, Vietnam

3 Development and Policies Research Center (DEPOCEN), Hanoi, Vietnam

4 School of Accounting, Auditing and Taxation, UNSW Sydney, Sydney, Australia

5 School of Business and Management, RMIT University Vietnam, Ho Chi Minh, Vietnam

1 Introduction and context

This paper is concerned with the intersection between two strands of the literature, namely, corruption and innovation. That innovation is a main engine for economic growth and development is widely accepted in the literature (Aghion et al. 2014; Aghion and Howitt 1998; Almeida and Fernandes 2008; Grossman and Helpman 1994; Romer 1990). At the micro level, innovation determines the growth of firms (Audretsch 1995; Audretsch et al. 2014; Brouwer et al. 1993; Coad and Rao 2008). Firm innovation[1] is shaped by both internal and external factors. Internal factors include ability of firms to innovate, firm investment, and research and development (R&D). External factors relate to the business environment, where prevalence of corruption has significant effects on firm innovation in both low-income countries (Ayyagari et al. 2014; de Waldemar 2012; Nguyen et al. 2016; Paunov 2016; Xie et al. 2019) and high-income countries (Dincer 2019; Ellis et al. 2019; Heo et al. 2020; Wen et al. 2020).

Corruption, generally defined as abuse of public office for personal gains (June et al. 2008: 6), is pervasive across nations and over time. It is widely agreed that corruption has significant effects on economic development (Kaufmann et al. 2011; Leff 1964; Mauro 1995; Tanzi and Davoodi 2002). These studies provide two opposing hypotheses about this phenomenon, namely, "sand-the-wheels" and "grease-the-wheels." The second, counterintuitive hypothesis is mainly advanced and supported in the presence of weak institutions (Aidt 2003; Méon and Sekkat 2005; Méon and Weill 2010).

According to the grease-the-wheels viewpoint, corruption may reduce the time spent in queues (Lui 1985), improve the quality of the civil service (Bayley 1966; Leys 1965), reduce the regulations of firm entry (Dreher and Gassebner 2013; Klapper et al. 2006) and promote efficient growth (Vial and Hanoteau 2010). Supporting the sand-the-wheels hypothesis, significantly negative effects of corruption on investment are confirmed by, for example, Mauro (1995) and Wei and Shleifer (2000). Negative effects of corruption on economic growth are transmitted via its effect on political stability, human capital, and private investment (Mo 2001) or on productivity (Lambsdorff 2003). Empirical studies suggest that whether corruption is sanding or greasing the wheels of firm innovation is largely shaped by contextual factors.

An important form of corruption in many developing countries is tax administration corruption. Yet, there exist to date very few, if any, rigorous studies that explicitly examine the impact of tax administration corruption on either innovation inputs or innovation outputs, possibly due to the lack of relevant data at the firm level. The present study attempts to fill the gap in the literature by providing an empirical analysis

[1] According to the Organisation for Economic Co-operation and Development (OECD) (2005), firms' innovative activities include both innovation inputs and outputs. The former refers to such activities as expenditure in R&D and investment in non-R&D activities to improve productivity or quality of output. The latter refers to the introduction of new products or new production processes, improvement of existing products, new marketing products, registered patents, etc.

of the impact of tax corruption on innovation of small and medium sized enterprises (SMEs) in Vietnam, using longitudinal data from biennial surveys of SMEs in Vietnam from 2005 to 2015.

Vietnam is chosen for the study for several reasons. First, Vietnam has a dynamic economy in transition with an increasing number of SMEs in the private sector. From a market perspective, Vietnam's continuing growth depends crucially on the growth and innovation of its private sector. Secondly, like many other developing countries, Vietnam has been suffering from widespread corruption in general and tax corruption in particular. In fact, corruption has been perceived by many different stakeholders as one of the most critical issues facing Vietnam at present (World Bank and Government Inspectorate of Vietnam 2012). Thirdly, and interestingly, despite widely reported tax corruption, Vietnam is classified as a high-tax-collection and high-tax-effort country in a study of determinants of tax level (Le et al. 2012: 25).

The present paper seeks to make three contributions to the existing literature on the effects of corruption on firm innovation in Vietnam. The first contribution is that the paper appears, to the best of our knowledge, to be the first that explicitly and rigorously investigates the impact of tax corruption on innovation inputs and outputs at the firm level, whereas Nguyen et al. (2016) is only concerned with general corruption on innovation outputs in Vietnam. Note that, in our study, tax corruption is confined to tax administration corruption, which is measured by the amount of tax bribe from business taxpayers to tax auditors, while Nguyen et al. (2016) only consider the propensity (yes or no) of general bribery. The second contribution is to provide theoretical arguments that shed lights on the positive effect of tax administration corruption on firms' innovative activities, including both innovation inputs and outputs. That is, our theoretical model concurs with the grease-the-wheels hypothesis, which tends to hold for transition or developing economies. The third contribution is the application of an appropriate empirical strategy to the problem under study, which was not employed in Nguyen et al. (2016). More specifically, we apply the control function method for dynamic nonlinear panel data models. This estimation methodology is employed to deal with endogeneity, unobserved heterogeneity, and initial condition problems simultaneously (Giles and Murtazashvili 2013; Michler and Josephson 2017; Papke and Wooldridge 2008; Wooldridge 2005). Our results show that tax corruption has a far greater impact on innovation than the 'naïve' random effects (RE) and the dynamic RE models.

The remainder of this paper is structured as follows. Section 2 provides a brief review of the literature whereas Sect. 3 briefly considers tax corruption and corporate taxation in Vietnam. Section 4 discusses the theoretical framework and estimation strategy, which is followed by sources of data and variable definitions in Sect. 5. In Sect. 6, the empirical results obtained are interpreted and discussed. Section 7 offers some concluding remarks and policy recommendations.

2 Literature review

In this section, we will first review theoretical arguments and then empirical studies. There are a few theoretical studies on the impact of corruption on innovation. From the perspective of organizational theory, Luo (2005) identifies how corruption negatively affects innovation via two main channels: bribery as a substitute to innovation, and interpersonal trust and trustworthiness. Using game theory, Veracierto (2008: 35) constructs a formal model showing that, under certain parameter ranges, better control of corruption results in an increase in the amount of product innovation. In contrast, it has been argued that corruption may increase efficiency and thus innovation via various channels (Leff 1964; Lui 1985; Mahagaonkar 2010). From this perspective, corruption speeds up the governmental process, reduces uncertainty, and introduces competition to scarce government resources. These two contrasting schools of thought correspond to the above-mentioned sand-the-wheels and grease-the-wheels, respectively.

The number of empirical studies linking corruption and innovation has been growing considerably in the past few years, partly due to data availability. First, there are single- and cross-country studies. Country-specific research focuses on developing or transition economies, for example, Bulgaria (Krastanova 2014), China (Trinh 2019; Xia et al. 2018; Xie et al. 2019; Xu and Yano 2017), Egypt and Tunisia (Goedhuys et al. 2016), India (de Waldemar 2012), Pakistan (Imran et al. 2020), and Vietnam (Nguyen et al. 2016). The only exception seems to be the US (Dincer 2019; Ellis et al. 2019). Cross-country studies also concentrate on emerging economies (Krammer 2019; Paunov 2016; Pirtea et al. 2019), African economies (Barasa 2018; Mahagaonkar 2010) and Eastern European and Central Asian economies (Habiyaremye and Raymond 2018; Kabadurmuş 2017). An exception is the member countries of the OECD (Wen et al. 2020) and a mixture of emerging and advanced nations (Heo et al. 2020).

Most empirical studies utilize firm-level data from various versions of the World Bank's Enterprise Surveys. Another important source of data is the European Bank for Reconstruction and Development (EBRD)'s Business Environment and Enterprise Performance Surveys (BEEPs). Some of these studies only consider cross-section of data (Krammer 2019) while others examine longitudinal data (Paunov 2016; Pirtea et al. 2019; Xu and Yano 2017). There are also a small number of studies which analyse panel data at the provincial level (Dincer 2019; Ellis et al 2019) or national level (Anokhin and Schulze 2009; Wen et al. 2020).

The studies that employ enterprise survey data measure innovation by the propensity to innovate (yes or no) and tend to focus on innovation outputs such a product or process innovation. An exception is Paunov (2016) who considers innovation inputs such as quality certificates and patents. Most of the firm-level studies measure corruption by using firms' information about spending on gifts or informal payments to public officials to get better services regarding customs, taxes, licenses, etc.

The studies that utilize more aggregated data tend to define innovation as innovation inputs, e.g., quantity and quality of patents (Dincer 2019) or resident patent applications and rates of realized innovation (Anokhin and Schulze 2009). Corruption at provincial level is measured in terms of number of corruption convictions or number of corruption stories reported in the press (Dincer 2019). At the national level, a popular measure of corruption is the control of corruption variable within the World Banks' World Governance Indicators.

Not surprisingly, empirical results concerning the broad effect of corruption on innovation are very heterogenous. Many studies have found that corruption has a negative impact on innovation (Anokhin and Schulze 2009; de Waldemar 2012; Dincer 2019; Habiyaremye and Raymond 2018; Paunov 2016). Nonetheless, other studies have concluded that corruption has a positive impact on innovation in the context of transition and developing economies (Ayyagari et al. 2014; Imran et al. 2020; Kabadurmuş 2017; Krastanova 2014; Nguyen et al. 2016, Xie et al. 2019).

The effects of corruption on innovation at a more detailed level are found to be dependent on various factors including types of innovation (corruption hinders product and organizational innovation but encourages marketing innovation; see Mahagaonkar 2010), types of corruption (Krammer 2019; Paunov 2016), firm size (smaller firms are badly affected; see Paunov 2016) or nature of the firm (state ownership and political connection matter; see Xu and Yano 2017). Wen et al. (2020) further suggest that is a threshold level of corruption control. Above this threshold further control is beneficial to innovation but below which it is not.

There are a few studies that relate to taxation in emerging economies. For instance, Sharma and Mitra (2015) show that, in India, tax-evading firms are likely to pay larger amounts of bribe to public officials than tax-compliant firms. Further, corruption is found to have a positive effect on product innovation but a negative effect on firms' efficiency. In the context of Chinese corporate tax reform and innovation, Cai et al. (2018: 2) argue that "lower taxes may reduce resources that firms spend on tax evasion, such as costs of bribing tax officers, which can be instead used on innovation activities."[2] The authors then find that lower corporate tax rate did stimulate innovation (R&D expenditure and number of patent applications) of medium and large enterprises as a result of the 2002 corporate tax cut in China.

The review of existing literature presented above reveals several research gaps. First, there is not yet a rigorous study that explicitly focuses on the direct effects of tax administration corruption on innovation in spite of the importance of the tax system in the reform and development of transition and developing countries (Hussain and Stern 1993). Secondly, while instrumental variables (IVs) have been employed to deal with endogeneity, insufficient attention has been paid to the simultaneous presence of endogeneity, state dependence and initial condition problem. The present paper attempts to address these gaps by examining the impact of tax corruption on innovative activities of SMEs in Vietnam, utilizing an appropriate estimation strategy.

3 Tax corruption and Vietnam's corporate taxation

3.1 What is tax corruption?

As a subset of general corruption, tax corruption refers to the unlawful exercise of public office by tax officials for their personal benefits (June et al. 2008: 12). The

[2] This seems to be a natural extension of a previously made argument that "firms under greater competition pressure are more motivated to avoid tax so as to have more investment money to compete in the market place" (Cai and Liu 2009: 765).

benefits to corrupted tax officials consist of not only financial but also non-financial gains. In the context of developing countries, tax corruption can be defined more accurately as “behavior on part of tax officials to improperly and unlawfully enrich themselves, or those close to them, by the misuse of the public power entrusted to them” (Li 1997: 475). This definition implies that in countries where the traditional culture encourages sharing, especially among members of an extended family including parents, spouses, children, and relatives or members of the same village (Vu et al. 2009), the well-being of people who are close to the corrupted tax officials is also of relevance.

Tax corruption, as defined above, could be classified in different ways depending on where the corruption happens in the operation of a tax system or its scale. In terms of the tax system operation, corruption could occur at the policy making, administration or dispute resolution stage. For example, corrupted tax policy makers could offer a tax incentive/exemption to a certain group of taxpayers that bribe them.[3] Similarly, a corrupted judge could biasedly and partially rule in favor of a taxpayer against the tax administrative agency (Riaz and Cantner 2020). In terms of scale, tax corruption can be grand or petty. As stated in the previous section, the scope of this paper is confined to tax administration corruption which tends to be recurrent and petty.

Tax corruption can take the form of bribery, embezzlement, theft, fraud, blackmail, extortion, collusion, and abuse of discretion (e.g., hiring unqualified family members or friends in tax departments). Tax corruption can involve one tax official, several tax officials, or between one or more tax officials and a taxpayer. An example of tax corruption that involves a tax official, and a taxpayer is the bribe that a taxpayer offers to a tax official in order to pass a tax audit or inspection. Tax corruption in this study is limited to the (illegal) interaction between taxpayers and tax officials.

As a result of tax corruption, the benefits to taxpayers typically involve less strict tax audits or lower tax liability. In some cases, taxpayers may also benefit from accessing privileged tax information that is beneficial to them. For tax officials, the benefits are ranging from cash, use or purchase of assets below market prices, payment for private expenses such as meals or traveling, recruitment/promotion of persons related to the tax officials in the taxpayer’s business.

In terms of process, the interaction between tax officials and taxpayers can be collusive or extortive. In collusive corruption, taxpayers and tax officials are involved in negotiating tax payment and amount of the bribery, and the process of negotiation can be explicit or implicit. In extortive corruption, tax officials initiate request of bribery in dealing with taxpayers. From the tax official’s perspective, there are widespread or loosely organized practices, i.e., a bribe sharing scheme between tax officials (Alm et al. 2016). In the remainder of this paper, unless otherwise stated, tax corruption refers to petty, recurrent briberies by taxpayers to tax officials.

[3] It is useful to draw a distinction between policy lobbying and policy corruption. Lobbying seeks to change the law, is transparent and does not involve direct benefits to politicians/public officials, whereas corruption tends to make an exception of the law, is secretive and involves direct benefits to the corrupted officials.

3.2 Corporate taxation and tax corruption in Vietnam

In Vietnam, an enterprise operates under one of the following five forms: (1) household enterprise, (2) private enterprise, (3) partnership, (4) limited liability company, and (5) joint stock company. These legal forms can be classified into two major categories of enterprises based on the tax payment method. The first category only includes household enterprises, which are not granted a tax code. The second one includes the remaining four forms of enterprises, which are registered with the local tax offices and granted a tax code.

The enterprises in the first category may pay a license tax, which is a lump-sum tax identified in the beginning of a fiscal year, value added tax (VAT), and/or personal income tax (PIT) depending on their sales revenue. They do not have to pay VAT and PIT if their annual turnover is less than 100 million VND, which is equivalent to about 4300 USD (at the February 2020 exchange rate). The local government tax officials will decide how much tax including license tax, VAT, and/or PIT an enterprise should pay in the current year based on the sales revenue of the enterprise in the previous year. In this situation, owners and managers of enterprises may personally deal with the tax officials about how much they have to pay, leading to possible tax corruption.

The enterprises in the second category have to calculate and self-declare all types of taxes they have to pay, including VAT and corporate income tax (CIT), and submit their queries to the online system of the General Department of Taxation (GDT) under the Ministry of Finance of Vietnam. Their monthly or quarterly submission depends on the size of their annual sales revenue. The local tax administration department will verify and randomly post-audit the tax filings of some enterprises for tax compliance. In addition, the local tax administration department conducts tax audit of all enterprises every 3–5 years. Tax corruption is likely to happen during visits of the local government tax officials to the enterprises.

As a transition economy, Vietnam is known for its bureaucratic administration and burdensome regulation. For example, in the 2015 calendar year, paying taxes in Vietnam took 540 h which was more than 2.5 times longer than the average of the East Asia and Pacific countries (198 h) (World Bank 2017).[4] Similarly, the number of tax payments in Vietnam in 2015 was 31 times which also far exceeded the average of the East Asia and Pacific Region (22.9 times). Facing this business/tax environment, there is an incentive for firms, especially innovating firms, to pay bribes/tax bribes to obtain better, faster and more certain government services and decisions (applications, licences, tax audits, etc.). In fact, many Vietnamese businesses perceive that corruption is a normal aspect of doing business and that they engage in corrupt activities to follow the 'rules of the game' (Nguyen et al. 2017: 305).

A 2012 survey sponsored by the World Bank and Government Inspectorate of Vietnam suggests that tax officials are identified by businesses as the public officials creating the most difficulties, and the ones that have been given the most unofficial

[4] This refers to the number of hours that a medium-sized company must spend to pay (or withhold) all taxes and mandatory contributions in a given year.

payments (i.e., briberies) and gifts (World Bank and Government Inspectorate of Vietnam 2012: 44–45). Nevertheless, unofficial payments are actively suggested by businesses (almost 90% of all cases), and only in about 10% of cases are the unofficial payments demanded (World Bank and Government Inspectorate of Vietnam 2012: 46). This has been confirmed by an independent survey based on a random sample of household businesses indicating that about 70% of the respondents always or often collude with tax auditors for mutual benefits (Nguyen et al. 2017: 305). Only a small fraction of respondents (about 13%) feel guilty about engaging in such an unlawful conduct.

4 The model and estimation strategy

4.1 The model

Based on the information provided in Sect. 3, we now seek to provide a discussion on the channels by which tax corruption can impact on innovative activities of SMEs in Vietnam. Our reasoning is similar but not identical to that of Cai et al. (2018), which has been discussed in Sect. 2.

Conceptually, the observed correlation between corruption and economic performance can result from a two-way causation, especially for transition and developing economies. A higher level of corruption may lower economic performance and, vice versa, a poorer economic performance may also encourage more corruption. For example, Ayyagari et al. (2014: 51) treat briberies as being dependent on innovation, and find that "innovators that pay bribes do not receive better services and do not have greater propensity to engage in other illegal activities such as tax evasion." Innovators are thus more likely to be victims than perpetrators of corruption.

The same kind of argument applies to tax administration corruption. The relationship between tax corruption and innovation can be intertwined and self-reinforcing, particularly in emerging economies. It is not implausible to argue that innovating firms tend to financially perform better than an average firm and can thus be targeted by corrupt tax officials who would demand larger amounts of tax bribery from them. This 'capacity to pay' argument may thus give rise to a reversed causality from innovation to tax corruption.

As mentioned in the previous section, while Vietnamese businesses often complain that tax officials use their prerogative and authority with a view to demanding more tax payments, bribes often result from a process of negotiation and collusion rather than extortion. This is because the accounting/tax rules and procedures in Vietnam are prescriptive with little room to manoeuvre. The tax officials have many discretionary powers, and taxpayers have no recourse to independent tax dispute resolution. The apparent conspiracy between business taxpayers and tax officials indicates businesses' desire to pass tax audits and also their darker motive of evading income taxation. As the most plausible result of negotiation, the amount of the bribe tends to be proportional to the amount of income tax that is in dispute, rather than to the business' profitability.

Tax corruption can assist firm-level innovation in Vietnam in two different ways. First, for SMEs, which often lack access to external funding, the tax 'savings' that results from tax corruption can represent an important source of funds. The tax savings, combined with funding from other sources, can be used to finance business expansion or improvement, including different types of innovation inputs such as R&D expenditure. Secondly, as 'tax-abiding' businesses, bribing firms would receive preferential treatments (relative to non-bribing firms) for any formal business applications including innovation.

Note that there are some subtle differences between our reasonings and those of Cai et al. (2018). Firstly, we argue that it is tax evasion that gives rise to tax savings that finance innovation whereas Cai et al. (2018:2) maintain that tax evasion takes resources away from innovative activities. Secondly, we focus on SMEs, which face constraints to raise funds externally, while Cai et al. (2018: 3) only consider medium and large enterprises. Thirdly and finally, in addition to the financial incentive, we also include the operational incentive for more certain and preferential treatments.

4.2 Estimation strategy

To examine the effect of tax corruption on innovation empirically, we augment the conventional, probit innovation model by incorporating tax bribery payment as an independent variable. We begin with discussing our estimation strategy to deal with state dependence and initial condition problems by the use of a dynamic binary response model. We then discuss how to apply the control function method for the dynamic binary response panel models in dealing with the endogeneity problem.

4.2.1 A dynamic nonlinear panel data model for state dependence

The RE probit model assumes zero correlation between unobserved individual effects and explanatory variables. Nevertheless, this assumption may be violated because, for example, current innovation of a firm may be a function of its past innovation, which is called the state dependence. Wooldridge (2005) proposes an approach to work with state dependence (unobserved heterogeneity) and initial conditions by using dynamic nonlinear panel data models. The state dependence and initial conditions in dynamic nonlinear panel data models may be derived from an underlying latent variable model as follows:

$$y^*_{1it} = \mathbf{z}_{1it}\beta_1 + \gamma y_{1it-1} + \alpha_{1i} + u_{1it}, y_{1it} = \mathbf{1}\left[y^*_{1it} \geq 0\right] \quad (i = 1, \dots, N; \quad t = 1, \dots, T) \tag{1}$$

$$y_{1i0} = 1\left[\mathbf{z}_{1io}\beta_0 + \alpha_{1i} + u_{1i0} \geq 0\right] \tag{2}$$

where y_{1it} is the observed dependent variable representing innovation of firm i at time t, y_{1i0} is firm i's innovation at the initial period, $\mathbf{z}_{1it}$ is a vector of strictly exogenous variables, and the asymptotic properties assume that T is small and fixed, and N is infinite. The error terms of the model have the distribution u_{1it} ~ N(0,1) whereas the time-invariant unobserved firm-effects (unobserved heterogeneity) are normally distributed, i.e., $\alpha_{1_i} \sim N(0, \sigma^2_{\alpha_1})$. The RE probit model assumes that unobserved heterogeneity is independent with explanatory variables ($\mathbf{z}_{1_{it}}$).

Another problem that needs to be addressed is the existence of initial conditions in estimating dynamic nonlinear panel data models. In other words, firm's innovation in initial period is correlated with α_{1_i}. For example, characteristics of firms and owners determine the decision to innovate in the initial period. To do so, Heckman (1981) provides the reduced form equation to solve the initial conditions problem, which is as follows:

$$y_{1i0} = 1\left[\mathbf{x}_{1i0}\beta_0 + v_{1i0} \geq 0\right] \tag{3}$$

$$v_{1i0} = \theta\alpha_{1i} + u_{1i0} \tag{4}$$

where $\mathbf{x}_{1_{i0}}$ is a vector, which can consist of the $\mathbf{z}_{1_{i0}}$ and/or exogenous instruments, and $v_{1_{i0}}$ is correlated with α_{1_i}, written as in Eq. (4). Note that $v_{1_{i0}}$ is uncorrelated with $u_{1_{it}}$ ($t \geq 1$). The log-likelihood functions for Eqs. (1) and (2) can be examined by using the Gaussian–Hermite quadrature.

In empirical studies, estimating the Heckman's estimator is time consuming when the density of y_{1i0} given ($\mathbf{x}_{1_{i0}}, \alpha_{1_i}$) is computed. Wooldridge (2005) uses the conditional maximum likelihood estimator, which treats the distribution conditional on the initial period value. This approach is similar to the strategy of Mundlak (1978) and Chamberlain (1984), which is called correlated RE approach. The form for α_{1_i} suggested by Wooldridge (2005) is

$$\alpha_{1i} = a_1 y_{1i0} + \overline{\mathbf{z}}_{1i} a + c_{1i} \tag{5}$$

where $\overline{\mathbf{z}}_{1_i} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_{1_{it}}$ is a vector of time averages of $\mathbf{z}_{1_{it}}$. Substituting it into Eq. (1) gives

$$y_{1it} = 1\left[\mathbf{z}_{1it}\beta_1 + \gamma y_{1it-1} + a_1 y_{1i0} + \overline{\mathbf{z}}_{1i} a + c_{1i} + u_{1it} \geq 0\right] \tag{6}$$

Equation (6) can then be estimated by the RE probit model, which solves the unobserved heterogeneity and initial conditions problems.

4.2.2 A dynamic nonlinear panel data model with endogenous explanatory variables

The approach by Wooldridge (2005) mentioned above only works if $\mathbf{z}_{1_{it}}$ is strictly exogenous, and it therefore cannot deal with unobserved heterogeneity and endogenous problems, which may exist simultaneously. Papke and Wooldridge (2008), and Giles and Murtazashvili (2013) suggest the application of the control function method in a setting of dynamic nonlinear panel data model to control for endogeneity of explanatory variables.

The tax corruption variable included in $\mathbf{z}_{1_{it}}$ in our study is, however, potentially endogenous as tax corruption decisions are correlated with the unobserved characteristics of firms and firm owners. Using the dynamic nonlinear panel data models with endogenous explanatory variables to estimate the relationship between tax corruption and innovation of SMEs in Vietnam is, thus, appropriate. Specifically, we rewrite Eq. (6) as

$$y_{1it} = 1\left[\mathbf{z}_{1it}\beta_1 + \gamma y_{1it-1} + \rho y_{2it} + \alpha_{1i} + u_{1it} \geq 0\right] \tag{7}$$

where $\mathbf{z}_{1it}$ is a vector of strictly exogenous variables, α_{1i} is an unobserved heterogeneity, u_{1it} is a serially uncorrelated idiosyncratic error term with $\mathrm{Var}\left(u_{1it}\right) = 1$, and y_{2it} is an endogenous variable.

First, we assume that the reduced form equation for the endogenous variable, y_{2it} is as follows:

$$y_{2it} = \mathbf{z}_{1it}\delta_1 + \mathbf{z}_{2it}\delta_2 + \alpha_{2i} + u_{2it} \tag{8}$$

where $\mathbf{z}_{2it}$ is a vector of IVs and u_{2it} are serially uncorrelated idiosyncratic error terms. Secondly, we assume that the unobserved heterogeneity in the first-stage equation, α_{2i}, is linear function of all exogenous variables, $\mathbf{z}_{it} = (\mathbf{z}_{1it}, \mathbf{z}_{2it})$ as

$$\alpha_{2i} = \overline{\mathbf{z}}_i\lambda + \eta_{2i} \tag{9}$$

where $\overline{\mathbf{z}}_i = \frac{1}{T}\sum_{t=1}^{T}\mathbf{z}_{it}$ is a vector of time averages of $\mathbf{z}_{it}$, and η_{2i} is error term. Equation (9) is consistent with the Mundlak's (1978) device for unobserved heterogeneity, α_{2i}.

Following Papke and Wooldridge (2008), we substitute Eqs. (9) in (8) and obtain:

$$y_{2it} = \mathbf{z}_{1it}\delta_1 + \mathbf{z}_{2it}\delta_2 + \overline{\mathbf{z}}_i\lambda + v_{2it} \tag{10}$$

where $v_{2it} = \eta_{2i} + u_{2it}$ is a new composite error term. We assume that $\left(u_{1it}, u_{2it}\right)$ has zero mean bivariate normal distribution and is uncorrelated to $\mathbf{z}_i$. This assumption implies that the error term in Eq. (7) is a function of the error term in Eq. (10).

$$u_{1it} = \theta u_{2it} + \xi_{1it} = \theta\left(v_{2it} - \eta_{2i}\right) + \xi_{1it} \tag{11}$$

where $\theta = \frac{\mathrm{Cov}\left(u_{1it}, u_{2it}\right)}{\mathrm{Var}(u_{2it})}$ with $\mathrm{Var}\left(u_{1it}\right) = 1$ and ξ_{1it} is a serially uncorrelated idiosyncratic error term. According to Giles and Murtazashvili (2013), the assumption in Eq. (11) is the contemporaneous endogeneity ofy_{2it}. If the contemporaneous v_{2it} explains sufficiently change of u_{1it} in Eq. (11), then y_{2it} become exogenous variable in Eq. (7).

Given our assumption above and Eq. (5),[5] we can rewrite Eq. (7) as follows:

$$\begin{aligned} y_{1it} &= 1\left[\mathbf{x}_{it}\beta + a_1 y_{1i0} + \overline{\mathbf{z}}_i a + c_{1i} + \theta\left(v_{2it} - \eta_{2i}\right) + \xi_{1it} \geq 0\right] \\ &= 1\left[\mathbf{x}_{it}\beta + a_1 y_{1i0} + \overline{\mathbf{z}}_i a + \alpha_{0i} + \theta v_{2it} + \xi_{1it} \geq 0\right] \end{aligned} \tag{12}$$

where $\mathbf{x}_{it} = (\mathbf{z}_{1it}, y_{1it-1}, y_{2it})$, $\beta = \left(\beta_1, \gamma, \rho\right)'$ is a vector of coefficients to be estimated, and $\alpha_{0i} = c_{1i} - \theta\eta_{2i} = \alpha_{1i} - \theta(v_{2it} - u_{2it})$ is composite unobserved heterogeneity. According to Giles and Murtazashvili (2013), and Michler and Josephson (2017),

[5] We have controlled for the initial conditions problems following Wooldridge (2005).

we must control for the relationship between α_{0i} and v_{2it} in distinctive time periods because it affects consistent estimates of the parameters from Eq. (12). Similar to our approach to the unobserved heterogeneity in the reduced form Eq. (9) for y_{2it}, we also assume that α_{0i} is independent of the initial conditions, y_{1i0} and the exogenous variables $\mathbf{z}_i$, but not of v_{2i}.

$$\alpha_{0i} = \lambda_0 \overline{v}_{2i} + \eta_{1i} \tag{13}$$

where $\overline{v}_{2i} = \frac{1}{T}\sum_{t=1}^{T} v_{2it}$ and η_{1i} is an error term, which uncorrelated to $\mathbf{z}_i$, y_{1i0}, and v_{2i} .

We now plug Eq. (13) into Eq. (12) and obtain:

$$y_{1it} = 1\left[\mathbf{x}_{it}\beta + a_1 y_{1i0} + \overline{\mathbf{z}}_i a + \theta v_{2it} + \lambda_0 \overline{v}_{2i} + \eta_{1i} + \xi_{1it} \geq 0\right] \tag{14}$$

Equation (14) solves the unobserved heterogeneity and endogenous problems, which occur simultaneously and initial conditions problems in dynamic nonlinear panel data models. In particularly, we follow the work of Giles and Murtazashvili (2013) with the two-step estimation procedure. First, we use a pooled OLS model to estimate Eq. (10) and obtain the residuals $\widehat{v}_{2it}$ from the reduced form equation and computed $\overline{\widehat{v}}_{2i} = \frac{1}{T}\sum_{t=1}^{T} \widehat{v}_{2it}$. Next, we employ the RE probit models to estimate Eq. (14). The standard errors in second stage were corrected by a bootstrap procedure because it obtains asymptotic standard errors for the estimation (Giles and Murtazashvili 2013; Papke and Wooldridge 2008). In recent years, Giles and Murtazashvili (2013), and Michler and Josephson (2017) have utilized this methodology to estimate the poverty dynamics in China and Ethiopia, respectively.

Following Svensson (2003), we control for the potential endogeneity of tax corruption by using IVs. He argues that firms have to pay bribes when dealing with public officials. In the dataset, we have information about the time that owners/managers of the surveyed SMEs spent working with government officials in dealing with regulations, which include tax issues. In addition, there is information about whether the SMEs have been inspected by government officials for various reasons including tax inspection. The possibility of having to pay tax bribes and the amount of tax bribe is likely to be higher when an SME is inspected by the government officials and its owner/manager spends more time working with them. Moreover, the local government tax officials randomly select enterprises to visit, and this process does not depend on firm performance as discussed in Sect. 3 that our IVs do not directly affect dependent variable. Therefore, we use two IVs to correct for the endogeneity of the tax corruption variable, which are (1) average % of owners/managers' working time spent each month dealing with government regulations and officials; and (2) the status of having been inspected by government officials for various reasons including tax inspection.

With the two IVs, the reduced form Eq. (10) becomes

$$\mathrm{TC}_{it} = \mathbf{z}_{it}\delta_1 + \mathrm{DealGov}_{it}\delta_{21} + \mathrm{Inspect}_{it}\delta_{22} + \overline{\mathbf{z}}_i\lambda + v_{2it} \tag{10'}$$

where TC_{it} is tax corruption defined as the ratio of firm's tax bribery payment to its total value added; $\mathrm{DealGov}_{it}$ and $\mathrm{Inspect}_{it}$ are two IVs, which determine TC_{it} but do not directly affect innovation of the SMEs; $\overline{\mathbf{z}}_i$ is a vector of time averages of the explanatory variables[6] in $\mathbf{z}_{it}$. We use the pooled OLS model to estimate Eq. (10′) and obtained the residuals $\hat{v}_{2it}$ from the reduced form equation and computed $\bar{\hat{v}}_{2i} = \frac{1}{T}\sum_{t=1}^{T}\hat{v}_{2it}$. We then plugged them into Eq. (14′) derived below to control for the endogeneity problems.

To analyze the effects of tax corruption on various types of innovation of SMEs, we apply the control function method to estimate the dynamic RE panel data model with endogenous explanatory variables. Equation (14) is written explicitly as follows:

$$\mathrm{IN}_{it} = 1\left[\mathbf{z}_{it}\beta + \gamma \mathrm{IN}_{it-1} + a_1\mathrm{IN}_{i0} + \rho \mathrm{TC}_{it} + \bar{\mathbf{z}}_i a + \theta v_{2it} + \lambda_0 \bar{v}_{2i} + \eta_{1i} + \xi_{1it} \geq 0\right] \tag{14′}$$

where IN_{it} is a dummy variable for innovation of firm i at time t. Innovation at time t is affected by innovation status at time $t-1$, which is indicated by IN_{it-1}, and innovation status in the initial period, which is indicated by IN_{i0}.[7] This approach is similar to what have been proposed by Giles and Murtazashvili (2013), Michler and Josephson (2017), Papke and Wooldridge (2008), and Wooldridge (2005). Incorporation of innovation status in the past in Eq. (14′) allows us to address potential correlation between unobserved firm heterogeneity and the other covariates.

5 Data and variables

The panel data used for this study is drawn from the biennial surveys of SMEs, which were conducted in Vietnam from 2005 to 2015. The surveys have been jointly conducted by the United Nations University World Institute for Development Economic Research (UNU-WIDER) in collaboration with the Economic Development Research Group at the University of Copenhagen, and the Central Institute of Economic Management (CIEM) and the Institute of Labor Science and Social Affairs (ILSSA) in Vietnam. In each round of survey, over 2500 SMEs across 12 industries in 10 provinces were randomly sampled.

One advantage of using this panel of data is that it contains information about various business aspects of the surveyed SMEs including their characteristics, production activities, and different types of innovation achieved by the enterprises. A wide range of monetary information about the bribe payments (for getting connected to public services; obtaining licenses and permits; dealing with taxes and tax collection; gaining government contracts/public procurement; dealing with customs/imports/exports) is also included in the data. We only extract the tax-bribe payment and the non-tax-bribe payment. From the tax-bribe payment, we are able to compute the cost of tax corruption (TC) as a percentage of value added of the SMEs, which is defined as the difference between total sales revenue and intermediate costs. As a result of data extraction, the total number of SMEs in our analysis data during the period of 2005–2015 adds up to 10,888.

[6] As we mentioned in Eq. (9), they are used to control for unobserved heterogeneity (Mundlak 1978).

[7] The initial period is defined as the first time that firms were observed in the dataset.

Table 1 presents definitions and descriptive statistics of the variables used in our estimation models. Major variables of interest include tax corruption and various indicators of innovation. The available data allows us to distinguish between incremental and more radical types of innovation outputs, which are assumed to be of different levels of technological difficulty. The former type is indicated by the improvement in existing products. The latter one is defined as either achievement of new products or new production processes. In terms of innovation inputs, we distinguish between machinery investments for innovation and R&D expenditure.[8] As a result, we have innovation inputs including machinery investments for innovation (MaInnovation) and R&D expenditure (R&D), and innovation outputs consisting of improved product (ImpProduct), new product or new production process (NewInnovation) and general innovation (Innovation), which is either improved product or new product or new production process.

Table 2 provides coefficients of the correlation matrix of the variables and their variance inflation factors (VIFs). As expected, the correlation coefficients are positive but there is no problem with multicollinearity because all VIF values are less than 10.

6 Results and discussion

6.1 Identification of tax corruption

We first perform the tests of validity of IVs in our model. The results of these statistical tests, summarized in Table 3, confirm the validity of the two IVs. In particular, the highly statistically significance of the Anderson's canonical correlation confirms the adequate explanatory power of our IVs for all categories of innovation. The Cragg–Donald–Wald F statistic and Stock–Wright LM statistic tests reject the null hypothesis of weak identification test and weak-instrument-robust inference for all types of innovation, respectively. The Sargan statistic tests do not reject the null hypothesis of over-identifying restrictions for all innovative activities. These results indicate that our instruments are statistically valid. Further, the endogeneity tests confirm the endogeneity of tax corruption for all innovative activities.

The reduced form regression is performed to estimate the determinants of tax bribery payment of the SMEs. The pooled-OLS estimation results in Table 4 suggest that among others our IVs, which are DealGov and Inspect, have significantly positive effects on the tax bribery payment of the SMEs. This finding indicates that our instruments are valid.

[8] This classification is based on the OECD's Oslo manual guidelines for collecting and interpreting innovation data (OECD 2005).

Table 1 Definitions and summary descriptive statistics of variables

Variable	Description	Mean	Sd	Min	Max
Dependent variables					
Innovation outputs					
ImpProduct	Having improved an existing product = 1; otherwise = 0	0.33	0.47	0.00	1.00
NewInnovation	Having introduced a new product or new production process = 1; otherwise = 0	0.20	0.40	0.00	1.00
Innovation	Having improved existing product, introduced new product, or introduced new production process = 1; otherwise = 0	0.40	0.49	0.00	1.00
Innovation inputs					
MaInnovation	Having investments in machinery to replace old equipment, improve productivity, improve quality of output, or produce a new output = 1; otherwise = 0	0.44	0.50	0.00	1.00
R&D	Having investments in R&D or purchasing patents = 1; otherwise = 0	0.01	0.09	0.00	1.00
Independent variable					
TC	Ratio of firm's tax bribery payment to its total value added	0.11	0.59	0.00	21.61
Control variables					
Formal	Belonging to the second legal form of SMEs, i.e., being either a private enterprise, partnership, limited liability company, or joint stock company = 1; Household business = 0	0.49	0.50	0.00	1.00
Export	Exporting products = 1; otherwise = 0	0.04	0.20	0.00	1.00
FirmSize	Natural logarithm of the total number of permanent workers	1.65	1.06	0.00	6.99
Training	Having provided some trainings to more than 50% of all workers = 1; otherwise = 0	0.12	0.32	0.00	1.00
QWorkers	Ratio of the workers who hold university or college degrees to the total number of permanent workers	0.03	0.06	0.00	1.00
University	Owners/managers of the SMEs having attained undergraduate or graduate degree = 1; otherwise = 0	0.17	0.38	0.00	1.00
Instrumental variables					
DealGov	SME owners/managers' time spent on dealing with government regulations (including tax regulations) and government officials (including tax officials) = 1; otherwise = 0	0.78	0.41	0.00	1.00
Inspect	Having been inspected by government officials for various reasons including tax inspection = 1; otherwise = 0	0.32	0.47	0.00	1.00

Table 2 Correlation matrix

	VIF	1	2	3	4	5	6	7	8	9	10	11	12
1. ImpProduct		1.000											
2. NewInnovation		0.328***	1.000										
3. Innovation		0.858***	0.606***	1.000									
4. MaInnovation		0.171***	0.188***	0.203***	1.000								
5. R&D		0.062***	0.070***	0.070***	0.091***	1.000							
6. TC	1.02	0.075***	0.073***	0.072***	0.065***	0.030**	1.000						
7. Formal	1.50	0.150***	0.102***	0.141***	0.011	0.044***	0.089***	1.000					
8. Export	1.17	0.091***	0.097***	0.101***	0.064***	0.070***	0.012	0.096***	1.000				
9. FirmSize	1.90	0.248***	0.207***	0.241***	0.198***	0.117***	0.083***	0.367***	0.345***	1.000			
10. Training	1.17	0.109***	0.121***	0.112***	0.078***	0.084***	0.028**	0.158***	0.149***	0.317***	1.000		
11. Qwokers	1.34	0.113***	0.125***	0.115***	0.074***	0.099***	0.074***	0.192***	0.151***	0.400***	0.192***	1.000	
12. University	1.42	0.094***	0.105***	0.096***	0.058***	0.078***	0.070***	0.179***	0.199***	0.432***	0.219***	0.401***	1.000

*$p<0.10$; **$p<0.05$; ***$p<0.01$

Table 3 Endogeneity tests of instrumental variables

	ImpProduct (1)	NewInnovation (2)	Innovation (3)	MaInnovation (4)	R&D (5)
Anderson canon. corr. (under-identification test)	37.74***	37.74***	37.74***	37.74***	37.74***
Weak identification test (Cragg–Donald Wald F statistic)	18.87***	18.87***	18.87***	18.87***	18.87***
Stock–Wright LM statistic Chi-sq(2)[a]	10.74***	14.04***	13.14***	24.97***	9.64***
Sargan statistic (p value)	0.011 (0.917)	1.135 (0.287)	0.019 (0.891)	1.099 (0.295)	2.694 (0.107)
Endogeneity test	9.077***	11.060***	11.553***	20.928***	6.036**

The results are derived from command ivreg2 in Stata

$*p<0.10$; $**p<0.05$; $***p<0.01$

[a]Weak-instrument-robust inference

Table 4 Reduced form estimation of tax corruption in the first stage

Dependent variable: tax corruption (TC)	
DealGov	0.071***
	(0.010)
Inspect	0.047***
	(0.018)
Formal	0.025
	(0.022)
Export	−0.084
	(0.058)
FirmSize	−0.008
	(0.020)
Training	−0.009
	(0.031)
QWorkers	0.051
	(0.165)
University	0.037*
	(0.022)
_Cons	0.066**
	(0.033)
Provincial location dummies	Yes
Year dummies	Yes
Industry dummies	Yes
N	10,888
R^2	0.026

Robust standard errors in parentheses (*$p<0.10$; **$p<0.05$; ***$p<0.01$). The industry dummies are for industries including food product; beverages; textiles; apparel and leather products; wood products; paper products; printing and reproduction of recorded media; petroleum products, chemical, pharmaceutical and plastics; non-metallic mineral products; basic metal products; electronic products, equipment, machinery, transport equipment; and furniture and others. The provincial location dummies include Ha Noi; Phu Tho; Ha Tay; Hai Phong; Nghe An; Quang Nam; Khanh Hoa; Lam Dong; Ho Chi Minh City; and Long An. Regressions include time averages of explanatory variables

6.2 Discussion of findings

We determine the effects of tax corruption on innovation outputs and innovation inputs in Eq. (14′) by applying the control function method for dynamic RE panel data models with endogenous explanatory variable. The second-stage regression results, reported in Tables 5 and 6, present average marginal effects of various

Table 5 Average marginal effects of determinants of innovation outputs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	RE			Dynamic RE			Dynamic RE with endogeneity		
	ImpProduct	NewInnovation	Innovation	ImpProduct	NewInnovation	Innovation	ImpProduct	NewInnovation	Innovation
TC	0.037***	0.018***	0.036***	0.025***	0.024*	0.023**	0.286*	0.529***	0.567***
	(0.008)	(0.005)	(0.009)	(0.010)	(0.013)	(0.011)	(0.167)	(0.097)	(0.199)
Lag of ImpProduct				0.052***			0.054***		
				(0.014)			(0.014)		
The initial ImpProduct				0.065***			0.064***		
				(0.013)			(0.012)		
Lag of NewInnovation					0.035***			0.034***	
					(0.013)			(0.009)	
The initial NewInnovation					0.029***			0.027***	
					(0.010)			(0.007)	
Lag of Innovation						0.069***			0.069***
						(0.015)			(0.017)
The initial Innovation						0.071***			0.070***
						(0.014)			(0.016)
Formal	0.038***	0.049***	0.060***	0.028	0.049***	0.043**	0.019	0.031***	0.024
	(0.011)	(0.008)	(0.011)	(0.018)	(0.015)	(0.019)	(0.022)	(0.011)	(0.024)
Export	0.041*	0.044**	0.080***	0.105***	0.058*	0.158***	0.126***	0.099***	0.201***
	(0.023)	(0.018)	(0.027)	(0.036)	(0.030)	(0.043)	(0.040)	(0.022)	(0.059)
FirmSize	0.061***	0.052***	0.070***	0.059***	0.037***	0.061***	0.061***	0.040***	0.064***
	(0.006)	(0.004)	(0.006)	(0.013)	(0.011)	(0.014)	(0.013)	(0.008)	(0.019)
Training	0.049***	0.050***	0.049***	0.029	0.038***	0.023	0.030	0.040***	0.026
	(0.014)	(0.011)	(0.015)	(0.019)	(0.014)	(0.021)	(0.022)	(0.010)	(0.020)

Table 5 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	RE			Dynamic RE			Dynamic RE with endogeneity		
	ImpProduct	NewInnovation	Innovation	ImpProduct	NewInnovation	Innovation	ImpProduct	NewInnovation	Innovation
QWorkers	0.198***	0.271***	0.268***	0.134	0.174*	0.160	0.119	0.141*	0.127
	(0.074)	(0.059)	(0.082)	(0.110)	(0.091)	(0.122)	(0.126)	(0.079)	(0.167)
University	0.014	0.003	0.004	−0.004	−0.004	−0.018	−0.013	−0.023**	−0.037
	(0.014)	(0.011)	(0.015)	(0.016)	(0.012)	(0.017)	(0.021)	(0.009)	(0.024)
v2							−0.260	−0.527***	−0.547***
							(0.167)	(0.097)	(0.196)
v2bar							−0.003	0.003	0.004
							(0.019)	(0.010)	(0.025)
Provincial location dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	10887	10888	10888	7449	7451	7451	7449	7451	7451

Columns (1)–(6) and Column (7)–(9) present robust standard errors and robust bootstrapped standard errors with 100 replications in parentheses, respectively; *$p<0.10$; **$p<0.05$; ***$p<0.01$. Other notes about industry and location dummies in Table 4 apply. Regressions include time averages of explanatory variables. The v2 and v2bar are the contemporaneous endogeneity (first-stage residuals free of serial correlation) and mean of the contemporaneous endogeneity, respectively

Table 6 Average marginal effects of determinants of innovation inputs

	(1) RE	(2) Dynamic RE	(3) Dynamic RE with endogeneity	(4) RE	(5) Dynamic RE	(6) Dynamic RE with endogeneity
	MaInnovation	MaInnovation	MaInnovation	R&D	R&D	R&D
TC	0.044***	0.039***	0.685***	0.001	0.003**	0.109**
	(0.010)	(0.013)	(0.205)	(0.001)	(0.001)	(0.047)
Lag of MaInnovation		0.054***	0.054***			
		(0.015)	(0.015)			
The initial MaInnovation		0.055***	0.055***			
		(0.014)	(0.015)			
Lag of R&D					0.001	0.001
					(0.006)	(0.006)
The initial R&D					0.005	0.004
					(0.006)	(0.006)
Formal	0.009	0.013	−0.009	0.002	0.002	−0.001
	(0.012)	(0.020)	(0.023)	(0.002)	(0.004)	(0.004)
Export	0.016	0.033	0.085	0.004	0.014**	0.022***
	(0.028)	(0.048)	(0.054)	(0.003)	(0.005)	(0.007)
FirmSize	0.125***	0.086***	0.089***	0.004***	0.006**	0.007**
	(0.006)	(0.015)	(0.017)	(0.001)	(0.003)	(0.003)
Training	0.071***	0.077***	0.081***	0.007***	0.004	0.005*
	(0.016)	(0.021)	(0.025)	(0.003)	(0.003)	(0.003)
QWorkers	0.173**	0.286**	0.248	0.033***	0.032**	0.027*
	(0.085)	(0.134)	(0.170)	(0.012)	(0.014)	(0.014)
University	−0.006	0.001	−0.022	0.002	0.001	−0.003
	(0.016)	(0.018)	(0.024)	(0.002)	(0.003)	(0.003)

Table 6 (continued)

	(1) RE	(2) Dynamic RE	(3) Dynamic RE with endogeneity	(4) RE	(5) Dynamic RE	(6) Dynamic RE with endogeneity
	MaInnovation	MaInnovation	MaInnovation	R&D	R&D	R&D
v2			−0.649***			−0.106**
			(0.204)			(0.047)
v2bar			0.004			-0.002
			(0.023)			(0.003)
Provincial location dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	10888	7451	7451	10549	7030	7030

Notes: Columns (1), (2), (4), and (5), and Column (3), and (6) present robust standard errors and robust bootstrapped standard errors with 100 replications in parentheses, respectively.* $p<0.10$; ** $p<0.05$; *** $p<0.01$. Other notes about industry and location dummies in Table 4 apply. Regressions include time averages of explanatory variables. The v2 and v2bar are the contemporaneous endogeneity (first-stage residuals free of serial correlation) and mean of the contemporaneous endogeneity, respectively factors on innovation outputs and innovation inputs, respectively.[9] Regression results in Table 5 for three types of innovation outputs including improvement of existing products (ImpProduct), introduction of new products or new processes (NewInnovation) and general innovation (Innovation). In the first three columns, we report the results of the naive RE panel data model with robust standard errors in parentheses. In the second set of three columns, we report the results of the dynamic RE panel data model without controlling for the endogeneity problem (with robust standard errors in parentheses). In the last three columns, we apply the control function method for the dynamic RE panel data model, controlling for the endogeneity problem and report the results with the bootstrapped-100-replication standard errors in parentheses. Regression results in Table 6 for two types of innovation inputs including machinery investments for innovation (MaInnovation) in the first three columns, and R&D expenditure (R&D) in the last three columns. The results from each group of three columns are the naive RE panel data model, the dynamic RE panel data model, and the control function method for the dynamic RE panel data model with endogeneity (dynamic RE with endogeneity), respectively. In these regressions, we include the time effects, industrial effects, provincial location effects, and time averages of control variables.

We focus on three main points: (1) the effect of tax corruption on innovation inputs and outputs, (2) the effect of past innovation on current innovation, and (3) the influence of other control variables on innovation. Regarding (1), the key finding of our analysis is that tax corruption has positive and statistically significant effects on all three types of innovation outputs for SMEs in Vietnam. Specifically, in Table 5, 1% point increase in the ratio of tax bribe payment to value added would increase the likelihood of improvement of existing products (Column 7), introduction of new products or new production processes (Column 8), and general innovation (Column 9) by about 0.29, 0.53, and 0.57% points, respectively. We provide robust analysis of innovation inputs by investigating whether or not tax corruption can represent an important source of funds as the tax 'saving.' It is shown that tax corruption has positive and statistically significant effects on all two types of innovation inputs for SMEs in Vietnam, which is similar to the case of innovation outputs. In particular, in Table 6, 1% point increase in the ratio of tax bribe payment to value added would increase the likelihood of machinery investments for innovation (Column 3), and R&D expenditure (Column 6) by about 0.69, and 0.11% points, respectively. In contrast, Paunov (2016) finds corruption (informal payments for obtaining licenses and permits) has a negative impact on machinery investments for innovation. Finally, these results from applying control function method for the dynamic RE model with endogenous variable show that tax corruption has more ten-fold to 30-fold size of effects on innovation than the naive RE and the dynamic RE models, depending on the type of innovative activities of the firm.

[9] From the reduced form regression based on Eq. (10'), we obtain the predicted values of the residuals $\widehat{v}_{2it}$ and $\widehat{\overline{v}}_{2i}$, which are then plugged into Eq. (14). This process is the control function method. We use the Stata xtprobit command to estimate Eq. (14') in the structural form. Finally, the average marginal effects are obtained by using margins command in Stata.

The key finding about the positive impact of tax corruption on innovation of SMEs in Vietnam confirms the grease-the-wheels hypothesis, which has also been supported in previous studies on the impact of general corruption on innovation (Krammer 2019; Nguyen et al. 2016; Sharma and Mitra 2015; Xie et al. 2019). Our finding is not implausible, particularly in the context of a transition economy in which the market mechanism, government regulation and tax administration are known to operate in an incomplete and inefficient manner. As discussed in Sect. 4, given the discretionary power enjoyed by tax auditors over taxpayers, briberies to tax officials could generate the funds that finance innovation, and the short-term certainties that taxpayers seek to conduct and expand their businesses.

Since tax corruption is known to be widely practiced in Vietnam, it seems reasonable to assume that Vietnamese SMEs may consider paying tax bribes as a normal way of doing business. Thus, they are willing to engage in tax corruption so long as the benefits (more certainties to conduct business and to innovate) exceed the costs (amount of bribery). In this sense, tax bribe payment can be seen as a means to facilitate new business opportunities, including innovation. From an economic perspective, tax bribery payment may be viewed as an instrument that supplements the government regulations.

Concerning (2), from Columns (4)–(9) in Table 5 and Column (2), and Column (3) of Table 6, we find consistent results that innovation outputs and only machinery investments of innovation inputs of an SME are positively determined by its innovation two years earlier and innovation in the initial period. The magnitude of the average marginal effect is, however, small. On average, an SME having obtained innovation two years earlier has a higher possibility of achieving current innovation by no more than seven percentage points compared to others. The same finding holds for an SME with innovation in the initial period. However, we do not find similar results for R&D expenditure, as shown in columns (5)–(6) in Table 6.

Finally, regarding (3), the results summarized in Tables 5 and 6 are economically plausible. More specifically, control variables such as firm size (FirmSize), exporting status (Export), employee training (Training), and proportion of professional employees (QWorkers) all have the expected positive impact on all types of SME innovation. This reassuring outcome provides a further evidence of the overall goodness of the estimated models.

6.3 Summary conclusions and policy implications

The present study attempts to shed light on tax corruption by empirically analyzing its effects on innovative activities of SMEs in Vietnam. While tax corruption can take many different forms, the focus of this study is on administration corruption that typically arises from the interaction between taxpayers and tax auditors during on-site tax inspection. To the best of the authors' knowledge, the study is the first that explicitly examines the influence of tax corruption on innovative activities of SMEs. A particular strength of the present study is that, unlike previous quantitative studies on tax corruption, the control function method for dynamic RE panel data models and IVs are employed to overcome problems associated with endogeneity of tax corruption, unobserved heterogeneity, and initial condition problems of innovation simultaneously.

Using a panel of data from surveys of SMEs in Vietnam, the study suggests that paying tax bribery facilitates all types of SME innovative activities, which include both innovation inputs and innovation outputs. The estimation results are statistically valid and robust against a number of diagnostic tests. The results obtained are also plausible in the sense that the control variables (such as firm size, exporting status, employee training, and proportion of professional employees) all have the expected positive impact on innovative activities of SMEs.

This key finding of the study supports the grease-the-wheels hypothesis of tax corruption. It is also consistent with those results obtained from previous empirical studies on the effects of general corruption on firm development, especially in the context of transition economies. Our main interpretation of this unconventional finding is as follows. In many transition and developing countries, the market mechanism, government regulation and tax administration often do not operate completely or efficiently. In such cases, tax bribes could potentially produce short-term certainties and tax savings which may be beneficial to some business activities including innovation. Tax bribery payments can, in this sense, be said to facilitate innovation of SMEs, at least in the short term.

The findings of the present study have adverse implications to both businesses and tax authorities in Vietnam. First, we must unequivocally stress that our findings do not necessarily mean tax corruption is beneficial to firms in in the long run. In fact, we argue that the hidden (from official accounting statements) and ongoing costs of tax corruption can damage firms' development in the long term. Firms' myopic view of benefits from paying tax bribes can diminish their long-term integrity and strategic capability. In particular, continuous illegal payments of tax and other briberies would be detrimental to improving staff morale and work practice.

Secondly, in the presence of widespread business engagement in tax briberies, fighting tax corruption in transition economies such as Vietnam is problematic. Without visible and sustained pressure from the private sector, the government has little incentive to combat tax corruption despite its commitments to anti-corruption policies. This is particularly true in the case of Vietnam where the economy has experienced high growth rate, noticeable poverty reduction, and sufficiently high tax to GDP ratio.

Finally, the causes of tax administration corruption in Vietnam are various and many of which lie beyond the control of the tax authority (see Nguyen et al. 2017). Nevertheless, the direct causes appear to be (1) high degree of discretionary power of tax auditors, and (2) regular visits of tax auditors to large number of businesses, including SMEs. Thus, to reduce the incidence of tax briberies, the government could tackle those two issues. For example, to address (1) the government could consider simplifying tax laws and tax procedures. Similarly, the government could also use digital technology such as automation, e-filing, etc., to reduce the face-to-face interaction between business taxpayers and tax auditors.

Funding The research is funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant Number 502.01-2018.03.

Declarations

Conflict of interest Each of the authors declares that he has no actual or potential conflict of interest including any financial, personal, or other relationships with other people or organizations that could inappropriately influence, or be perceived to influence, his work.

Research involving human participants The research utilizes secondary data and does not involve human participants.

Informed consent Not applicable because there are no human participants.

References

Aghion P, Howitt P (1998) Endogenous growth theory. MIT Press, Cambridge
Aghion P, Akcigit U, Howitt P (2014) What do we learn from Schumpeterian growth theory? In: Aghion P, Durlauf S (eds) Handbook of economic growth, vol 2. Elsevier, Amsterdam, pp 515–563
Aidt T (2003) Economic analysis of corruption: a survey. Econ J 113(491):F632–F652
Alm J, Martinez-Vazquez J, McClellan C (2016) Corruption and firm tax evasion. J Econ Behav Organ 124(C):146–163
Almeida R, Fernandes A (2008) Openness and technological innovations in developing countries: evidence from firm-level surveys. J Dev Stud 44(5):701–727
Anokhin S, Schulze WS (2009) Entrepreneurship, innovation, and corruption. J Bus Ventur 24(5):465–476
Audretsch D (1995) Innovation, growth and survival. Int J Ind Organ 13(4):441–457
Audretsch D, Coad A, Segarra A (2014) Firm growth and innovation. Small Bus Econ 43(4):743–749
Ayyagari M, Demirgüç-Kunt A, Maksimovic V (2014) Bribe payments and innovation in developing countries: are innovating firms disproportionately affected? J Financ Quant Anal 49(1):51–75
Barasa L (2018) Corruption, transaction costs, and innovation in Africa. Afr J Sci Technol Innov Dev 10(7):811–821
Bayley D (1966) The effects of corruption in a developing nation. West Polit Q 19(5):719–732
Brouwer E, Kleinknecht A, Reijnen JO (1993) Employment growth and innovation at the firm level. J Evol Econ 3(2):153–159
Cai H, Liu Q (2009) Competition and corporate tax avoidance: evidence from Chinese industrial firms. Econ J 119(537):764–795
Cai J, Chen Y, Wang X (2018) The impact of corporate taxes on firm innovation: evidence from the corporate tax collection reform in China. In: National Bureau of economic research working paper series, no. 25146. https://doi.org/10.3386/w25146
Chamberlain G (1984) Panel data. In: Griliches Z, Intriligator M (eds) Handbook of econometrics, vol 2. Elsevier, Amsterdam, pp 1247–1318
Coad A, Rao R (2008) Innovation and firm growth in high-tech sectors: a quantile regression approach. Res Policy 37(4):633–648
de Waldemar F (2012) New products and corruption: evidence from Indian firms. Dev Econ 50(3):268–284
Dincer O (2019) Does corruption slow down innovation? Evidence from a cointegarted pnel of U.S. states. Eur J Polit Econ 56(1):1–10
Dreher A, Gassebner M (2013) Greasing the wheels? The impact of regulations and corruption on firm entry. Public Choice 155(3):413–432
Ellis J, Smith J, White R (2019) Corruption and corporate innovation. J Financ Quant Anal. https://doi.org/10.1017/S0022109019000735

Giles J, Murtazashvili I (2013) A control function approach to estimating dynamic probit models with endogenous regressors. J Econom Methods 2(1):69–87
Goedhuys M, Mohnen P, Taha T (2016) Corruption, innovation and firm growth: firm-level evidence from Egypt and Tunisia. Eurasian Bus Rev 6(3):299–322
Grossman G, Helpman E (1994) Endogenous innovation in the theory of growth. J Econ Perspect 8(1):23–44
Habiyaremye A, Raymond W (2018) How do foreign firms' corruption practices affect innovation performance in host countries? Industry-level evidence from transition economies. Innovation 20(1):18–41
Heckman JJ (1981) The incidental parameters problem and the problem of initial conditions in estimating a discrete time−discrete data stochastic process and some Monte Carlo evidence. In: Manski CF, McFadden DI (eds) Structural analysis of discrete data with econometric applications. MIT Press, Cambridge, pp 179–195
Heo Y, Hou F, Park SG (2020) Does corruption grease or sand the wheels of investment or innovation? Different effects in advanced and emerging economies. Appl Econ. https://doi.org/10.1080/00036846.2020.1791313
Hussain A, Stern N (1993) The role of the state, ownership and taxation in transitional economies. Econ Transit 1(1):61–87
Imran SM, Rehman HU, Khan REA (2020) Effect of corruption on firm level innovation: evidence from Pakistan. Econ Bus Lett 9(1):41–47
June R, Chowdhury A, Heller N, Werve J (2008) A user's guide to measuring corruption. UNDP Oslo Governance Centre, Oslo
Kabadurmuş FNK (2017) Corruption and innovation: the case of EECA countries. J Entrep Innov Manag 6(2):51–71
Kaufmann D, Kraay A, Mastruzzi M (2011) The worldwide governance indicators: methodology and analytical issues. Hague J Rule Law 3(2):220–246
Klapper L, Laeven L, Rajan R (2006) Entry regulation as a barrier to entrepreneurship. J Financ Econ 82(3):591–629
Krammer SMS (2019) Greasing the wheels of change: bribery, institutions, and new product introduction in emerging markets. J Manag 45(5):1889–1926
Krastanova P (2014) Greasing the wheels of innovation: how corruption and informal practices of firms impact the level of innovation in Bulgaria? Masters degree thesis, Central European University
Lambsdorff J (2003) How corruption affects productivity. Kyklos 56(4):457–474
Le T, Moreno-Dodson B, Bayraktar N (2012) Tax capacity and tax effort: extended cross-country analysis from 1994 to 2009. Policy Research Working Paper WPS6252. World Bank International Trade and Investment Unit, Washington, DC
Leff N (1964) Economic development through bureaucratic corruption. Am Behav Sci 8:8–14
Leys C (1965) What is the problem about corruption? J Mod Afr Stud 3(2):215–230
Li J (1997) Counteracting corruption in tax administration in transitional economies: a case study of China. Bull Int Fisc Doc 51(11):474–492
Lui F (1985) An equilibrium queuing model of bribery. J Polit Econ 93(4):760–781
Luo Y (2005) An organization perspective of corruption. Manag Organ Rev 1(1):119–154
Mahagaonkar P (2010) Money and ideas: four studies on finance, innovation and the business life cycle. International Studies in Entreprenuership 25. Springer, New York
Mauro P (1995) Corruption and growth. Quart J Econ 110(3):681–712
Méon P-G, Sekkat K (2005) Does corruption grease or sand the wheels of growth? Public Choice 122(1):69–97
Méon P-G, Weill L (2010) Is corruption an efficient grease? World Dev 38(3):244–259
Michler J, Josephson A (2017) To specialize or diversify: agricultural diversity and poverty dynamics in Ethiopia. World Dev 89(C):214–226
Mo P (2001) Corruption and economic growth. J Comp Econ 29(1):66–79
Mundlak Y (1978) On the pooling of time series and cross section data. Econometrica 46(1):69–85
Nguyen A, Doan H, Nguyen M, Tran-Nam B (2016) The impact of petty corruption on firm innovation in Vietnam. Crime Law Soc Chang 65(4–5):377–394
Nguyen N, Doan H, Tran-Nam B (2017) Tax corruption and private sector development in Vietnam. eJ Tax Res 15(2):290–311
Organisation for Economic Co-operation and Development (OECD) (2005) Guidelines for collecting and interpreting innovation data. OECD, Paris

Papke L, Wooldridge J (2008) Panel data methods for fractional response variables with an application to test pass rates. J Econom 145(1–2):121–133
Paunov C (2016) Corruption's asymmetric impacts on firm innovation. J Dev Econ 118(1):216–231
Pirtea MG, Sipos GL, Ionescu A (2019) Does courrupiton affect busines innovation? Insights from emerging countries. J Bus Econ Manag 20(4):715–733
Riaz MF, Cantner U (2020) Revisiting the relationship between corruption and innovation in developing and emerging countries. Crime Law Soc Chang 73(5):395–416
Romer P (1990) Endogenous technological change. J Polit Econ 98(5):S71–S102
Sharma C, Mitra A (2015) Corruption, governance and firm performance: evidence from Indian enterprises. J Policy Model 37(5):835–851
Svensson J (2003) Who must pay bribes and how much? Evidence from a cross section of firms. Quart J Econ 118(1):207–230
Tanzi V, Davoodi H (2002) Corruption, growth, and public finances. In: Jain A (ed) The political economy of corruption. Routledge, London, pp 89–100
Trinh M (2019) "Greasing" or "sanding-the-wheels"? An empirical assessment of the impact of corruption on firm innovation in China. Master's thesis, Lund University
Veracierto M (2008) Corruption and innovation. Econ Perspect 32(1):29–39
Vial V, Hanoteau J (2010) Corruption, manufacturing plant growth, and the Asian paradox: Indonesian evidence. World Dev 38(5):693–705
Vu HN, Sonobe T, Otsuka K (2009) An inquiry into the transformation process of village-based industrial clusters: the case of an iron and steel cluster in northern Vietnam. J Comp Econ 37(4):568–581
Wei S-J, Shleifer A (2000) Local corruption and global capital flows. Brook Pap Econ Act 31(2):303–354
Wen J, Wang QL, Feng G-F, Chen SW, Chang C-P (2020) Corrupion and innovation: linear and nonlinear investigation of OECD countries. Singap Econ Rev 65(1):103–129
Wooldridge J (2005) Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. J Appl Economet 20(1):39–54
World Bank (2017) Paying taxes, Doing business. World Bank, Washington, DC
World Bank and Government Inspectorate of Vietnam (2012) Corruption from the perspectives of citizens, firms and public officials: results of sociological surveys. National Political Publishing House, Hanoi
Xia H, Tan Q, Bai J (2018) Corruption and technological innovation in private small-medium scale companies: does female top management pay a role? Sustainability 10(7):1–18
Xie X, Qi G, Zhu KS (2019) Corruption and new product innovation: examining firms' ethical dilemmas in transition economies. J Bus Ethics 160(1):107–125
Xu G, Yano G (2017) How does anti-corruption affect corporate innivation? Evidence from recent anticorruption efforts in China. J Comp Econ 45(2017):498–519