Energy demand and factor substitution in Vietnam: evidence from two recent enterprise surveys

Phu Viet Le*

*Correspondence:
phu.le@berkeley.edu
Fulbright University Vietnam,
Ho Chi Minh City, Vietnam

Abstract

Vietnam's economy is one of the most energy-intensive economies in the world, facilitated by long-standing government policies indirectly subsidizing energy prices through various state-owned enterprises in the energy sector. A consequence of this is that firms are using too much energy in production. This raises a crucial issue as to whether Vietnam can continue its development trajectory in the new era with rising energy prices and increased awareness of the use of fossil fuels and environmental pollution. In this context, understanding energy use patterns and firms' behaviors regarding cheap energy prices is critical to forming appropriate energy policies and management practices. Using large-scale firm-level data, we have found explicit evidence of firms' substitution of energy for capital inputs. This effect is present in both the short term and long term and in many energy-intensive industries. These results indicate that there is substantial benefit in appropriate pricing of primary energy and electricity, while also providing credit incentives for capital investment in more energy-efficient equipment. Reducing the rate of growth of energy demand, averaging approximately 10% annually, will have significant macroeconomic impacts. A quicker transition to less energy-intensive economic growth will also help to protect both the environment and public health.

Keywords: Energy demand, KLEM, Elasticity of substitution, Translog production

1 Introduction

Over the past two decades, Vietnam has witnessed dramatic increases in energy demand. The current Vietnamese economy is extremely inefficient in terms of energy consumption. Electricity consumption has increased by approximately 13% per year, requiring a huge increase in generation capacity from a modest amount of 8.7 GW in 1990 to 27 GW in 2000 and more than 48 GW in 2018. Primary energy consumption more than doubled during 2005–2017, from 31 Mtoe (million tons of oil equivalent) to more than 75 Mtoe in 2017 (Fig. 1). After 30 years of economic reforms, Vietnam has become one of the most highly electrified countries in the world, with grid electricity reaching 98% of the population. Universal electrification and strong economic growth are driving the per capita consumption of electricity. Prior to the economic reform in 1986, Vietnam consumed only 70 kwh of electricity per person. This figure has now reached almost

Fig. 1 Primary energy consumption in Vietnam (source: Statistica (2019))
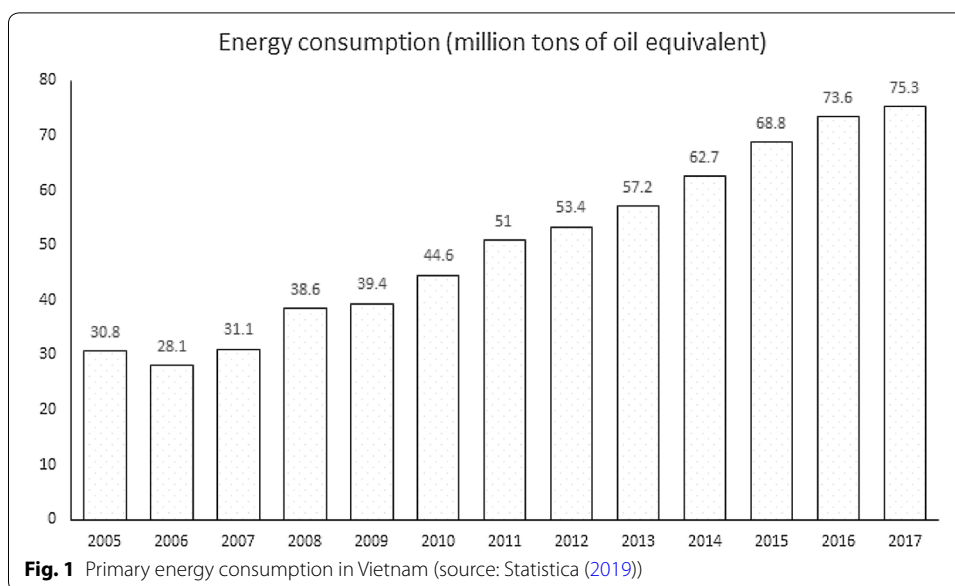
2000 kwh, despite the average per capita consumption still being much lower than that of many countries in the region (Malaysia 4600, Thailand 2500, and China 3900) (Dapice and Le 2018).

The elasticity of electricity to GDP (growth rate of electricity consumption/GDP growth rate) is a clear indicator of the high energy intensity in Vietnam. The elasticity in Vietnam is astoundingly high relative to other comparable countries over more than a decade of development, reaching 1.5–2.0 times the GDP growth rate (FPT 2015). It is well understood that, during the initial stage of economic development, the growth rate of electricity is often higher than is the rate of economic growth due to the development of energy-intensive industries. When the process of basic industrialization is complete, the economic structure changes, moving gradually to the development of low-energy but high-efficiency industries, and economic growth induces reductions in the energy intensity of GDP (Burke and Csereklyei 2016). Figure 2 shows the amount of primary energy (MJ) per GDP (PPP). At present, Vietnam's energy intensity is 5.94, slightly lower than in China (6.69), whereas it is much higher than in other regional countries in the ASEAN block [such as Malaysia (4.68), Indonesia (3.53), the Philippines (3.12)] and India (4.73). This means that Vietnam is using much more energy for a unit of economic output. If the projection is correct, in 2025, the energy intensity will be even 40% higher than it is today (Dapice and Le 2018).

The rapid increase in demand for energy is attributed to expanding industrial and construction sectors. For example, electricity consumption has grown with an annual average of almost 16%, from 9.1 TWh in 2000 to 69 TWh in 2014—seven times in size, and more than a half of the total electricity demand (FPT 2015). The four largest industries, paper, cement, steel, and fertilizer, together consumed 15% of the aggregate demand in 2010 (United Nations Development Programme 2011). This strong growth may be, in part, due to low energy prices and substitution of primary fuels for other inputs such as labor and capital. Both primary energy and electricity power for industrial and civil

Fig. 2 Primary energy intensity (MJ) per GDP (PPP) in 2015 (source: World Bank (2019))
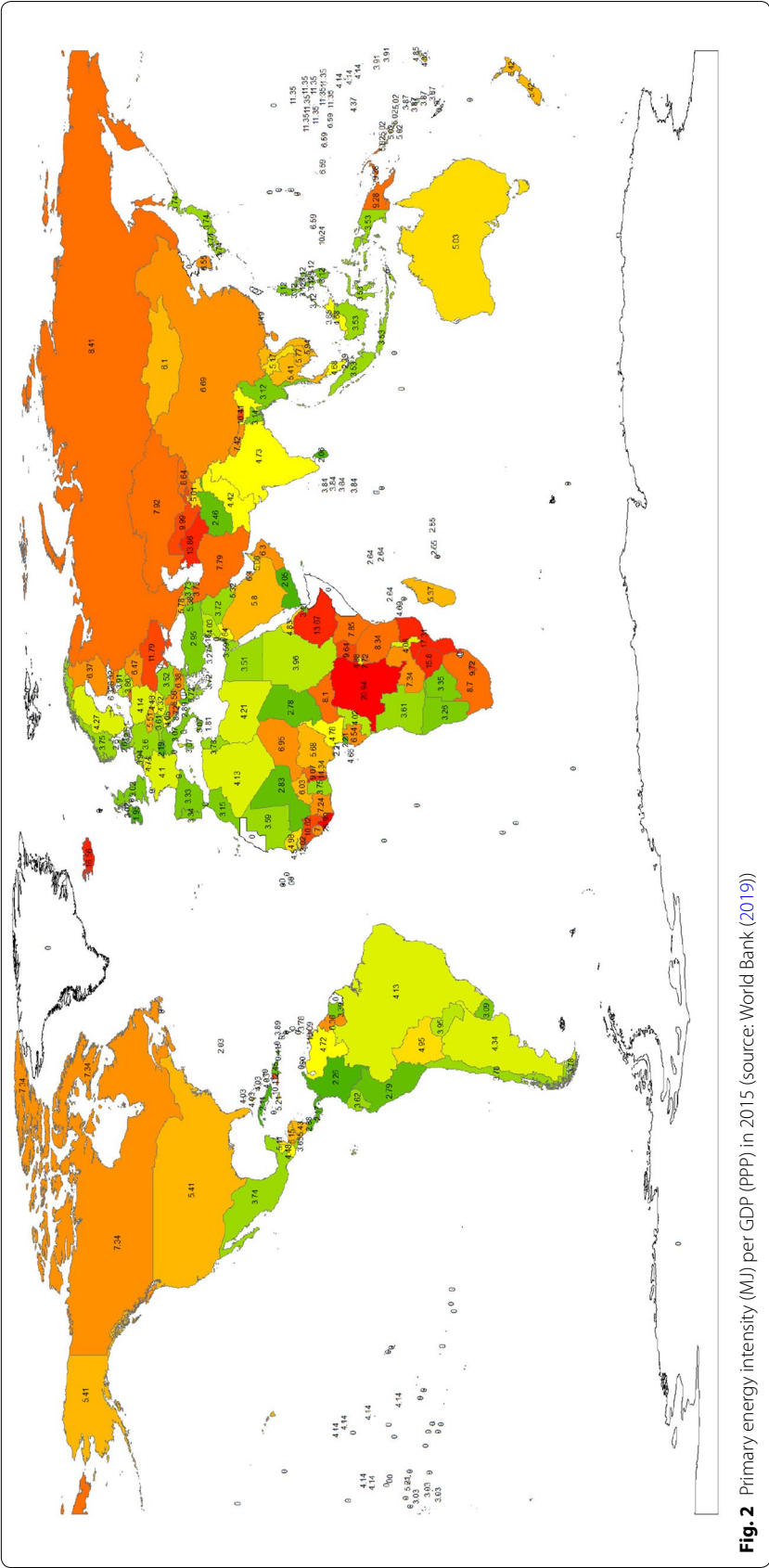
sectors in Vietnam are relatively cheaper than are those in other countries in the region and the world (Australian Energy Council 2016)—the legacy of the exploitation of hydropower, which has very low operating costs, as the main energy source and large implicit subsidies of fossil fuel extraction and consumption. In recent years, the coal price has increased by 40%, forcing customers to shift from fossil fuel to using more electricity (United Nations Development Programme 2014).

The retail prices of electricity and other primary energy are regulated strictly by the Vietnamese government. Electricity Vietnam (EVN), the wholesale buyer of electricity from independent power producers, is allowed to adjust the price only in small steps. The low retail price does not help either demand side management or energy efficiency efforts. Also due to a low retail price, the feed-in-tariff of renewable energy has been effectively low, taking into account weak power purchase contracts. As a result, the contribution of renewable energy (excluding hydropower, which has essentially been exhausted) to the total energy production is minimal at the moment, despite there being huge potential with both solar and wind power (Ministry of Investment and Trade 2017). It is clear that the Vietnamese government cannot maintain a policy of cheap energy prices to sustain economic development.

Therefore, understanding the characteristics of energy demand is an extremely important condition for designing appropriate energy and industrial development policies. This study is the first to investigate the nature of energy demand from the manufacturing aspect. It provides empirical evidence for the high energy intensity in the manufacturing sector in Vietnam. We have utilized two recent enterprise surveys, which collected information of all establishments in Vietnam, to estimate the production function, with four aggregate inputs in a KLEM—Capital (K), Labor (L), Energy (E), and Intermediate Material (M)—framework. We have found clear evidence of substituting energy for capital, taking advantage of the low energy cost in Vietnam. The evidence is widespread across many sectors, business types, and ownerships. Decomposing the effect by sector, we found that heavy industries, such as mining and manufacturing, more likely use energy as a substitute for capital. The effects are heterogeneous among firms, with small firms, privately owned, and foreign-invested firms tending to substitute energy for capital. This result has important policy implications, as Vietnam is struggling to meet a rising demand in the face of essentially exhausting all cheap energy supplies, and is experiencing environmental damage from increasing reliance on coal power.

2 Methodology and data

In production economics, a firm's behavior can be framed as either a profit maximization or a cost minimization problem, which gives rise to duality of the solutions. That is, a firm either chooses the optimal level of output, at a given total cost, to maximize the profit, or chooses the optimal levels of inputs to minimize the total cost of production, at a given level of output. The derived demands for inputs in a four-input KLEM model are determined by the level of production, the production technology, and the relative prices of inputs.

A key question in estimating either a production function or the derived demand for inputs is the degree to which inputs can be either substitutes or complements of each other. For example, some studies suggest that energy use is a complement to capital

(equipment, machinery) and a substitute for labor (Griffin and Gregory 1976), whereas others indicate that substitution does occur but to a limited extent (Berndt and Wood 1975). Establishing the property of the derived demand for inputs (either complement or substitute, and degree of complementarity and substitutability) is crucial to public policy and energy development strategy. If factors are highly substitutable, increasing the price of one input might serve to reduce its own demand while increasing the demand for the substitute. However, if inputs are either not close substitutes or are even complementary, then changing the price of one input might have unintended consequences. For example, if energy and capital are distant substitutes, as suggested by Berndt and Wood (1975), energy pricing policies that raise the cost of electricity might negatively affect production, as firms are not able to adjust by switching to other inputs.

The degree of either complementarity or substitutability is conveyed in the concept of elasticity. The elasticity, calculated as the percentage changes in quantity to a one percent change in price, indicates how easy it is for a firm to change an input demand in response to an external price shock. If the price of an input such as energy increases, it is expected that the firm will respond by reducing the use of energy. This response is governed by the substitution effect and the income effect. If energy is becoming relatively more expensive, firms will switch to other less expensive inputs, depending on either the substitutability or the technologies that allow for such an adjustment. The income effect is related to the cost share in the total cost. If an input is used in a small amount, a higher price would not affect the total cost by much, and, therefore, the income effect is expected to be low. However, if an input accounts for a large share of the total cost, the income effect would be high, and the firm is expected to respond more decisively to even a small price change. This is particularly relevant in energy pricing policy, because the cost share is often small. The average cost shares were approximately 5% for K, 27.5% for L, 4.5% for E, and 63% for M, on average, for the United States during 1947–1971 (Berndt and Wood 1975). Therefore, the responsiveness of the demand, and, thus, the effectiveness of energy pricing policy, depends largely on technological ability.

When there are multiple inputs in the production, the literature does not agree on either the pairwise complementarity or the substitutability between inputs. This was summarized in Chung (1987), where, for example, K and E were found to be either complementarities or substitutes in manufacturing in the United States and Canada, depending on the time duration (short-term or long-term), methodologies and assumptions, and data aggregation level (sector or country level). A short-term production function assumes that the capital input does not change, whereas a long-term model requires the capital stock to change to reflect the impact of capacity expansion or substitution between capital stocks and variable inputs. As a result, short-term functions are expected to show that variable inputs, such as L, E, and M, are substitutes, and K and E are complements. In the long term, K and E could be substitutes when firms are given sufficient time to invest in more energy-efficient equipment if the price of energy is expected to increase. The time duration is also expected to have a strong impact on the degree of substitutability, with a higher elasticity expected in the long term than in the short term (Uri 1979).

The demand for energy is a derived demand for end-use services rendered by combining capital equipment and energy inputs. Therefore, the long-term and short-term

effects are influenced by the duration for which a firm can replace capital goods. Hartman (1979) divided the decision-making involving residential or industrial energy demand into three levels. The first level is the decision whether to either buy or replace fuel-burning capital goods. The second level is the decision whether to buy equipment of certain technical and economic characteristics. The third level is to decide the frequency and intensity of use. In the short term, the stock of capital goods is fixed; thus, firms are only allowed to change the level of variable inputs. In the long term, a firm can change both capital goods as well as input intensity. In Hartman's framework, the short term corresponds to the third level, whereas the long term corresponds to the first and second levels. Typically, equipment used in heavy industries often last much longer than those in light industries or services sectors; therefore, complete adjustments in response to energy prices may take a lot of time. For example, the economic life of equipment in mining and construction business could be as long as 20 years (FAO 1992), while office equipment are replaced every 3–5 years.

Different modeling techniques produce different versions of the elasticity. Cross-sectional studies assume that the market is in equilibrium and that between-firm variations in the data reflect a long-term decision to utilize the observed patterns of inputs. In those models, the capital stock would adjust instantaneously to a change, or the expectation of changes, in the relative price of inputs. Accordingly, a cross-sectional study is assumed to produce the long-term elasticity. Meanwhile, models using repeated observations over a short period of time (such as time series or panel data) will identify the short-term elasticity because the capital stock normally does not change within the observed time. A firm may only change the level of variable inputs such as energy, materials, and labor, to some extent. Either modeling approach, whether cross-sectional or panel data, may encounter certain issues. A cross-sectional model is prone to omitted-variable bias, which may occur when a factor that influences the demand for energy is not included in the model (Greer 2012, Ch. 9). If the omitted factor has a direct impact on a firm's performance, then the estimated elasticity may be biased. Consequently, cross-sectional studies may produce an overly sensitive estimate of the price elasticity (Hartman 1979). With panel data, a first-difference estimator could solve the omitted-variable bias arising from an unobserved factor that does not change over time. However, the disadvantage is that there are often little variations in energy prices in a short panel, leading to large standard errors and low accuracy. Estimation based on longer time series may also be problematic if there is a structural change over time (Greer 2012, Ch. 9).

Most of the aforementioned studies used either state-level or sectoral data, which aggregate inputs across firms and geographic locations. An example is that of Griffin and Gregory (1976), which relied on cross-country variations in energy price to estimate the long-term energy demand. Depending on the aggregation level of data, energy demand can be estimated for countries, regions, sectors, or firms. The estimated elasticity is affected by aggregation levels. The finer the aggregation level, the lower the ability to substitute one input for another—which is due to specific technology and the production process being employed. Sectoral studies, such as Dargay (1983), found that the elasticity of energy demand differs substantially between Swedish manufacturing sectors—which could be a result of the distinct production structure of each industry. Dargay shows that energy and capital are

complementary in most industries, whereas not only energy and labor but also energy and intermediate goods are substitutes. This demand pattern indicates that raising the energy price could result in firms substituting away from energy, at the expense of investment in capital goods.

With the availability of micro-data at the firm level, a firm-level production function can be estimated (Kleijweg et al. 1990; Pitt 1985). The firm size is an important determinant of the energy demand. In Dutch manufacturing, small firms adjust more quickly to energy price changes than do large firms, reflecting the greater likelihood of large-scale capital-intensive production being fixed in the short term. It is also possible to decompose each input into different sub-categories, such as different types of energy used (electricity, coal, gasoline, diesel, and LPG). Then, within each input category, inter-fuel substitution can be observed to a greater extent (Pindyck 1979).

However, researchers have to balance between the aggregation level of the data and modeling complexity. Modeling the demand of specific fuel types is often complicated by the presence of corner solutions (Woodland 1993). Most firms do not use all types of fuel. Electricity and gasoline are most often used for operations and transportation. The use of fuels, such as coal, as a heat source is needed only in certain manufacturing process, such as steel and cement production. As a result, estimating the demand function for a specific fuel often runs into a corner solution; that is only some type of fuel is used in the production. The corner solutions arise from two possibilities: one is that some technology only requires certain inputs, and the other is that, given the relative prices of inputs, it is not economical to use all types of fuels. In econometric terms, the choice of fuel type is endogenous to prices and other factors. There are two remedies to this potential corner solution issue. One is to estimate a demand system for each sector separately, as in Woodland (1993). Another is to use the aggregate data-assuming implicit substitution between sub-types of fuel, as in most other studies.

2.1 Methodology

The most popular approach used in the literature is the transcendental (translog) model, which is an approximation of any twice-differentiable strictly quasi-concave homothetic production function with constant return to scale (Thompson 2006). It is assumed that any technical change affecting the aggregate inputs is Hicks-neutral (for example, changes in input quality may shift the balance of input ratios toward the use of better inputs). A KLEM function linking a firm's output to the four aggregate inputs—assuming symmetry and constant return to scale, and a stochastic term, is specified as:

$$
\begin{aligned}
lnQ_{it} = {} & a_0 + a_K lnK_{it} + a_L lnL_{it} + a_E lnE_{it} + a_M lnM_{it} \\
& + \frac{1}{2}\left[b_{KK}lnK_{it}^2 + b_{LL}lnL_{it}^2 + b_{EE}lnE_{it}^2 + b_{MM}lnM_{it}^2\right] \\
& + b_{KL}lnK_{it}lnL_{it} + b_{KE}lnK_{it}lnE_{it} + b_{KM}lnKlnM_{it} \\
& + b_{LE}lnL_{it}lnE_{it} + b_{LM}lnL_{it}lnM_{it} + b_{EM}lnE_{it}lnM_{it} \\
& + \sum_{j}\beta_j X_{it}^j + \gamma\,\mathrm{Year}_t + \sigma_i + v_{it},
\end{aligned}
\tag{1}
$$

where *it* represents firm i at year t, X is a set of variables to control for firm's characteristics, and σ_i is firm i's fixed effects, assumed to be unchanged during the observation time.

Symmetries of cross derivatives imply that $b_{KL} = b_{LK}$, $b_{KE} = b_{EK}$, $b_{KM} = b_{MK}$, $b_{LE} = b_{EL}$, $b_{LM} = b_{ML}$, and $b_{EM} = b_{ME}$. Furthermore, the homothetic condition requires that:

$$a_K + a_L + a_E + a_M = 1 \tag{2}$$

and

$$\begin{cases} b_{KK} + b_{KL} + b_{LE} + b_{KM} = 0 \\ b_{LK} + b_{LL} + b_{LE} + b_{LM} = 0 \\ b_{EK} + b_{EL} + b_{EE} + b_{EM} = 0 \\ b_{MK} + b_{ML} + b_{ME} + b_{MM} = 0 \end{cases} \tag{3}$$

Assuming that energy is paid with its marginal product, $Q_E = e$, with Q_E being the first-order derivative of the production function with respect to E, then the elasticity of output to energy input is also the cost share of energy, θ_E, and is derived as:

$$\theta_E = \frac{\partial lnQ}{\partial lnE} = a_E + b_{EE}lnE + b_{KE}lnK + b_{LE}lnL + b_{ME}lnM \tag{4}$$

In production economics, Allen's relative substitution elasticity (RSE) is an important concept. It measures the responsiveness of relative input uses to relative input prices. Firms change the input mix when relative prices change. The sign and speed of changes would indicate the type of inputs (complements or substitutes), and the degree of interdependence. Without explaining the intermediate steps, we have the general formula of the RSE for two factors Capital and Energy, under a constant return to scale assumption, as follows:

$$\alpha_{EK} = \frac{\%\Delta E/K}{\%\Delta r/e} = \frac{K * Q_K + L * Q_L + E * Q_E + M * Q_M}{K * E} \cdot \frac{|H_{EK}|}{|H|} = \frac{Q}{E * K} \cdot \frac{|H_{EK}|}{|H|}, \tag{5}$$

where $|H|$ is the determinant of the bordered Hessian matrix H of first- and second-order derivatives of the production function. H_{EK} is the EK cofactor of H.

$$H = \begin{bmatrix} 0 & Q_K & Q_L & Q_E & Q_M \\ Q_K & Q_{KK} & Q_{KL} & Q_{KE} & Q_{KM} \\ Q_L & Q_{LK} & Q_{LL} & Q_{LE} & Q_{LM} \\ Q_E & Q_{EK} & Q_{EL} & Q_{EE} & Q_{EM} \\ Q_M & Q_{MK} & Q_{ML} & Q_{ME} & Q_{MM} \end{bmatrix} \tag{6}$$

The elements in matrix H are derived as follows:

$$Q_{EE} = Q\frac{b_{EE} - \theta_E + \theta_E^2}{E^2}$$
$$Q_{KE} = Q\frac{b_{KE} + \theta_K\theta_E}{KE} \tag{7}$$
$$Q_E = \theta_E * \frac{Q}{E}$$

Once the substitution elasticity has been identified, the cross-price elasticity of energy demand to capital price is inferred directly from Allen's RSE:

$$\varepsilon_{EK} = \frac{d(lnE)}{d(lnr)} = \theta_K * \alpha_{EK} \tag{8}$$

These elasticities are calculated at given cost shares, typically at either the sample means or representative firms. The derivation of the mathematical formula and applications are detailed in Thompson (1997, 2006).

Estimation methods

The translog production function (1) can be estimated by either a single-equation or a system of equations approach. In a system of equations approach, the cost shares of all inputs, namely K, L, E, and M, will be estimated. Due to the singularity of covariance matrix of four equations (the cost shares of all inputs adding up to one), it is necessary to drop one equation prior to the estimation. In both approaches, the symmetries and homothetic constraints in (2–3) are imposed on the parameters of either the single Eq. (1) or the system of Eq. (9).

$$\begin{cases} \theta_K = a_K + b_{KK}lnK + b_{KL}lnL + b_{KE}lnE + b_{KM}lnM \\ \theta_L = a_L + b_{LK}lnK + b_{LL}lnL + b_{LE}lnE + b_{LM}lnM \\ \theta_E = a_E + b_{EK}lnK + b_{EL}lnL + b_{EE}lnE + b_{EM}lnM \\ \theta_M = a_M + b_{MK}lnK + b_{ML}lnL + b_{ME}lnE + b_{MM}lnM \end{cases} \tag{9}$$

In this study, we have employed the single-equation approach. Due to the availability of panel data, observed in 2015 and 2016, the advantage of fixed-effect and random-effect estimators outweighs the efficiency gained by the system approach using three-stage regression. Panel data with fixed effects can address a potentially serious concern about unobserved firm-specific characteristics that affect both the use of inputs and firm's performance. Theoretically, a 3SLS approach could potentially improve the estimation properties compared to the single equation (Thompson 2006); however, it does not take advantage of having repeated observations to control for omitted-variable bias.

We have presented five different methods to check the sensitivity of the results and shown the robustness of the choice of the single-equation approach. First, we estimated the translog production function in (1) with pooled data over the years without the homothetic constraints. Then, we estimated two models using panel data with fixed and random effects, in case there may be omitted factors that cause bias in the least-squares estimate. Third, we estimated the translog model with pooled data, with the homothetic constraints imposed on the parameters, which produces the long-term estimate of the elasticities. Finally, we applied a first-difference estimator to identify the short-term elasticities. To account for provincial difference, such as either the business environment or local policies, which may affect overall firm performance, we used standard errors clustered at the provincial level to adjust for correlation among firms in the same province.

To calculate the Allen's elasticities of substitution and price elasticities, all elements of the bordered Hessian matrix *H*, and of the cofactors must be identified, and then their determinants, which must be semi-negative definite, must be calculated. These were examined thoroughly in the calculation of the elasticities.

2.2 Data

We used two recent Vietnam Enterprise Surveys (VES), 2015 and 2016. The annual survey collects information from all formally registered establishments, enterprises, and cooperatives, either operational or idling, during the preceding year of the survey. The totals in 2015 and 2016 were 415,656 and 455,296 firms, respectively. The survey covers information relating to the ownership, industrial sectors (up to five digits by Vietnam's System of Economic Branches in 2007), locations, import/export activities, economic performance in the preceding year, the total number of employees and the total labor cost, assets and liabilities (including short- and long-term assets), tax and payables, R&D expenditures, and branches.

Most importantly, the surveys collected detailed information on which types of energy were used, the amount, and the total value. However, due to the survey design, only up to a quarter of the dataset contained information about energy consumption. To create a panel dataset from the two surveys, we matched tax identification numbers between the two individual datasets. There is an issue with sub-branches of the same companies that share the same tax code. It is difficult to separate individual branches' performances and the characteristics of firms having the same tax code at different locations and sectors. As a consequence, we dropped firms belonging to the same branch and sharing a unique tax code from the analysis. We further restricted the sample to include only firms that were operational in the previous year. The final sample, after dropping firms that reported a negative revenue, a negative capital stock, and a negative input cost (labor, energy, or material), and were excessively large, includes 149,959 observations, with 72,499 observations in 2015 and 77,460 observations in 2016. The balanced panel dataset has 37,685 firms.

The most common energy category is electricity usage, which includes the amount consumed (in 1000kwh), self-production, sales, net production, and the value of purchase. Other types of energy, including gasoline, coal, diesel, mazut oil, liquefied petroleum gas (LPG), and liquefied natural gas (LNG), are grouped into four main categories: gasoline, coal, diesel, and LPG. As expected, gasoline is the second most common source of energy with almost 43,000 observations, followed by diesel, with 28,000 observations. Other primary energy sources are used either very sparsely or to a limited extent. In the model, we aggregated the value of all energy types into a single energy category (E), assuming perfect substitution between the source and purpose of uses (heating/cooling or electric equipment operation). Though inter-fuel substitution is possible, modeling that possibility requires specific industry information on energy use, which is not available. Details of energy types and consumption are presented in Additional file 1: Table S1.

Other variables in the model include the total revenue of production or services as the dependent variable (Q); the total value of long-term assets (K), which includes long-term receivables and values of fixed assets at the end of the reporting year, less depreciation; the total cost of labor (L); and the total value of short-term assets, intermediates, and inventories as the material cost (M). The use of the value (in monetary terms) instead of the quantity (in units) in the production function assumes that the value is directly proportional to the quantity, after controlling for either sectoral differences or types of firm ownership. This assumption is entirely justifiable because most firms in Vietnam are small; thus, prices are taken as exogenous. Furthermore, the market of inputs, including

capital, labor, and energy, is well regulated in Vietnam. Energy prices, in particular, are strictly under government control. Thus, all firms are expected to face the same price schedule (although sectoral or ownership differences are allowable). Establishments are classified by 13 types of ownership (Additional file 1: Table S2), and exclude household business. We collapsed the five-digit standard industrial classifications into 21 major industrial sectors. However, there were no observations in two areas related to the communist party (VSIC = 84) and international cooperation (VSIC = 99). As a result, the sample contained 19 industrial sectors (Additional file 1: Table S3). The processed data are summarized in Additional file 1: Table S4.

3 Results and discussion

Table 1 shows the long-term cost shares (or output elasticity with respect to each input), the Allen's RSE, and the price elasticities, calculated at the sample means, from the estimated coefficients of the translog function with the homothetic constraints. Similar to Berndt and Wood (1975)'s estimates for the United States, the cost shares are small for energy and capital inputs, and large for labor and material inputs. The energy or capital expenses as a share of the total cost account for 3–10% of the total cost, with a higher share observed in traditionally energy-intensive sectors, such as energy, manufacturing, mining, services, and transportation.

The most interesting findings come from the Allen's elasticities of input substitution and the price elasticities. All own price elasticities are negative, as expected, with the most elastic demand belonging to energy (−1.69), indicating the possibilities of firms being able to adjust energy demand easily; whereas, the demand of capital, labor, and material is normally less flexible, thus having inelastic own price elasticity coefficients. The reason may be that labor normally signs a long-term contract, thus making it difficult to adjust on a regular basis.

The cross-price elasticities show that there are possibilities of substitutes and complements among all four inputs. Focusing on energy, $\varepsilon_{EK} = .0033$, $\varepsilon_{EL} = -1.3572$, $\varepsilon_{EM} = -.047$ suggest that capital and energy are used as substitutes (increasing the price of capital will cause firms to shift to using more energy, at a given output), whereas energy and labor, and energy and material are used as complements (increasing either the labor cost or the material cost will reduce the demand for energy). This may be due to variable inputs, such as labor, energy, and materials, being used often in a certain proportion. However, decomposing the elasticities by sectors, firm sizes, and ownerships has revealed widely different patterns of cost shares and substitutions, a phenomenon

Table 1 Estimated cost shares, Allen's elasticity of substitution, and price elasticities at sample means (long-term effects)

	Cost share	Allen's RSE				Price elasticities			
	θ	K	L	E	M	K	L	E	M
K	0.0313	− 17.0617				− 0.5341			
L	0.4453	0.7941	− 1.7644			0.0249	− 0.7857		
E	0.0657	0.1058	− 3.0478	− 25.6660		0.0033	− 1.3572	− 1.6875	
M	0.4576	− 1.9246	1.3333	− 0.7145	− 1.5318	− 0.0602	0.5938	− 0.0470	− 0.7010

also observed in many developed and developing countries during their initial development phase (Woodland 1993). The findings also agree with Dargay (1983) that estimates based on aggregating data could be misleading due to the heterogeneity between sectors and firms.

Based on the estimated positive cross-price elasticity of substitution, there is evidence that energy–capital substitution has occurred in the agriculture, energy, manufacturing, mining, services, transportation, water and sanitation, and other sectors (Additional file 1: Table S5). These are also the largest energy users in the economy, accounting for more than 60% of the total electricity demand (FPT 2015). Substitution also occurred between variable inputs, notably energy and material, in administration, communication, construction, finance, entertainment and sciences sectors. Those sectors have relatively low energy cost shares. However, energy and labor are shown to be complements in all sectors. Substitution could occur only if the production processes allow greater use of one input, supposedly less expensive, to replace more expensive inputs. For example, firms could invest in a long-term asset (i.e., productive capital, such as more energy-efficient equipment) to reduce the short-term demand for variable inputs such as energy. Explaining why either substitutions or complements occurred in certain sectors requires detailed characterization of the production function, which is beyond the scope of this study. The full list of the RSE and price elasticities is provided in Additional file 1: Table S6.

Firm ownership may also influence the motivation to substitute energy for capital, as observed in private firms (whether fully privately owned or less than 50% government owned), foreign-owned firms, and other types of joint ventures (Additional file 1: Table S7). These firms are more profit-driven than are other types of firm ownerships; thus, using more energy could help lower the total cost of production. This finding raises a potentially serious issue: whether foreign-directed investment (FDI) is flocking to Vietnam to take advantage of low energy prices. There has been a persistent impression about Vietnam's attempts to lure foreign investment at all costs. One of the most recognized incentives is that the Vietnamese government has been offering very low energy prices, either of electricity or primary fuels, relative to other regional countries. It is unclear whether it is the low energy price, or other factors, that explain the attractiveness of the business environment. Garg et al. (2015) surveyed 4,000 registered FDI projects and found that the energy cost share accounts for less than 5% of the total operating cost and that respondents are willing to pay for better energy infrastructure and service quality. However, Tang et al. (2016) found a causal relation between energy consumption and economic growth, implying that either a higher energy price or other measures to conserve energy consumption would likely hamper long-term development.

It is important to note that the estimated elasticities in Additional file 1: Table S5 or Additional file 1: Table S7 are derived for either the sectoral or the subsample averages. All firms within either the same sector or subsample are assumed to have the same cost structure. However, this may not necessarily be the case. Therefore, we estimated the price elasticities among firms of different sizes. First, we partitioned the dataset into 10 deciles, each approximately one-tenth of the sample size, according to the size of the total revenue, and then estimated the cross-elasticities within each decile (Additional file 1: Table S8). We found strong and consistent substitution effects between capital and

energy in small- and medium-sized firms, up to either an average total revenue of VND 16.614 bn (almost USD 780 thousand) or an average labor force of 56 people. Interestingly, large firms do not exhibit capital–energy substitution. This finding is critically important in electricity pricing policy. The data clearly indicate pervasive use of electricity in production, while other primary fuels are used rarely. This may be a consequence of both a low price of electricity and limited ability to invest in equipment that is more energy efficient.

Sensitivity checks

Five different models were estimated, including an unconstrained pooled regression (model 1), unconstrained fixed effects and random effects panel regressions (models 2–3), and pooled and first-differencing regressions with the homothetic constraints (models 4–5). The full regression results are provided in Additional file 1: Table S9. The use of panel data with fixed effects was to address a potential omitted-variable bias in the least-squares method. For example, unobserved industrial knowledge of the production or the climate condition might affect industrial productivity and the use of energy for either heating or cooling purposes. Having an unobservable input that correlates with both the dependent variable (production output) and the explanatory variables (input expenditures) will cause biases in the estimated coefficients. If these unobserved characteristics remain constant during the study period of 2 years, then they could be discarded by a first-difference estimator. We used the estimated coefficients in Additional file 1: Table S9 to calculate the cost shares in Additional file 1: Table S10.

The cost shares are largely similar between the unconstrained pooled regression (column 1, Additional file 1: Table S10), the random-effects regression (column 3), and the constrained regression (column 4). The fixed-effects (column 2) and first-differencing (column 5) models show a lower energy share and a higher capital share. These require some explanation. First, it is only meaningful to compare the constrained model using the full data (model 4) with the constrained model using the first-differences (model 5). It is challenging to estimate a constrained model in the presence of panel data. To incorporate the homothetic constraints in the production function, data will need to be transformed into first-differences. However, due to the imbalanced nature of the data, only 37,647 firms were used in the first-differencing model. Second, the interpretation of the model using pooled data and the model using first-differences is different. The pooled regression, utilizing cross-sectional variations between firms, is consistent with the long-term interpretation of the elasticity; whereas, the model using year-to-year variations (or first-differencing data) will identify the short-term effect. We have still found clear evidence of substitution of energy for capital in the short term (Table 2).

4 Conclusion and implications for energy policy

It is critical for the Vietnamese government to acknowledge the extensive pattern of energy substitution for capital in many energy-intensive industries. Given the extremely high energy intensity of the economy, a low energy price, and vast untapped renewable energy potential, there is a lot of room for improvements on both the demand and supply sides of the energy market. We have listed three areas that have received much attention recently. The first is increased energy efficiency through effective pricing reforms. A

Table 2 Estimated cost shares, Allen's elasticity of substitution, and price elasticities at sample means (short-term effects)

	Cost share	Allen's RSE				Price elasticities			
	θ	K	L	E	M	K	L	E	M
K	0.1370	− 10.4106				− 1.4259			
L	0.4898	− 0.9815	− 1.3888			− 0.1344	− 0.6802		
E	0.0399	0.1827	− 2.5166	− 29.0308		0.0250	− 1.2326	− 1.1574	
M	0.3334	− 2.8132	1.3361	0.3006	− 3.0827	− 0.3853	0.6544	0.0120	− 1.0277

more appropriate energy price will reduce the incentives of firms to substitute energy for capital and encourage investments in more energy-efficient capital. Pricing of primary energy and electricity should be done on a competitive basis, absent of either direct or indirect government interventions.

The current average retail price of electricity in Vietnam is low, about 8.1 US cent/kwh, and is among the lowest in Southeast Asia and the world. There is a significant cross-subsidy scheme that guarantees a very low price for manufacturing and administrative sectors (6.1 c/kwh for normal hours, 3.8 c/kwh for off-peak hours, and 11 c/kwh for peak hours) by charging a higher price for residential consumption (from 7.2 to 12.6 c/kwh, on an increasing block rate) and even a much higher price for the services sector (11.44 c/kwh for normal hours, 6.96 c/kwh for off-peak hours, and 19.69 c/kwh for peak hours). In contrast, the average price is about 10 c/kwh in Indonesia, 8–12 c/kwh in India, 10 c/kwh in China, 11 c/kwh in Malaysia, 30 c/kwh in the Philippines, and is even much higher in developed countries, such as 22–47 c/kwh in Australia and 31 c/kwh in Germany (Dapice and Le 2018). In fact, the average retail price of electricity in Vietnam is kept below the long-term marginal cost of production at 9.37 c/kwh (ADB 2016).

The low price is attributed, in part, to government policies subsidizing the energy sector. Most fossil fuel subsidies in Vietnam are indirect, provided to state-owned enterprises in energy production and distribution by means of lax environmental regulations, low interest credits, and low-cost access to labor, land, and primary inputs such as coal and petroleum products, leading to forgone state tax and increasing debt (United Nations Development Programme 2014). For example, the indirect government subsidy of the electricity sector amounted to USD 2.86bn out of the total fossil fuel consumption subsidy of USD 3.45bn in 2012, based on a well-established price-gap approach of the International Energy Agency (IEA). Concerning environmental impacts, the current environmental protection tax is US 17 c/liter of gasoline and US 8.6 c/liter of diesel. Meanwhile, a ton of lignite coal only pays USD 1.29, effectively making the coal tax less than one percent of that of gasoline per unit of emissions discharge. These subsidy schemes and preferential treatment of coal have distorted the downstream energy market and failed to achieve an environmentally just allocation of resources.

Second, there is a great potential in improving energy efficiency through technical regulations and enforcements. For example, in the commercial building sector, up to 40% of consumption could be conserved through improvements in air conditioning, auxiliary equipment (water pump, blower), lighting system, office equipment, and elevator systems. In the cement sector, this number could be up to 50% (GIZ 2017).

Of course, the prerequisite condition for a successful policy roll-out is an appropriate energy pricing reform and provision of financial support for equipment purchases and retrofits, given that most Vietnamese enterprises are small to medium companies with limited resources to incur large capital investments.

Third, in the long-term, the government must accelerate the restructuring of the economy to shift away from energy-intensive industries to high value-added and less energy demanding sectors, such as services and tourism. With the right policies in place, Vietnam could accelerate a clean energy transition by capitalizing on the renewable energy evolution. Despite significant private sector interest in solar and wind power, foreign investors have been reluctant to invest due to cumbersome land acquisition, the unbankability of the power purchase agreement, and various unofficial costs. More transparency and power market liberalization will alleviate the roadblock.

Specifically, the Vietnamese government should gradually deregulate the electricity market to allow independent power producers and retailers to compete in all segments of the market, from generation, to transmission and retailing. Developing clean power from solar radiation and wind should be among the top priorities, considering vast locally untapped potential of these resources. Fully internalizing environmental costs into the cost of production will help technologies with high financial costs but more environmentally friendly become more competitive. Removing implicit subsidies for fossil fuels will help bring the cost of production closer to the actual cost of generation. Only targeted subsidies should be retained where they can be justified as serving social welfare objectives, such as maintaining a lifeline tariff for the poor (Dapice and Le 2018). A new initiative, the direct power purchase agreement (DPPA), which matches the supply from independent renewable energy producers with the demand from large buyers, is currently being explored. At the moment, Electricity Vietnam, a state-owned enterprise, has the sole control over the national transmission grid, and over 60% of the system's total generation capacity. There are possibilities for foreign investors to take a minor stake in transmission upgrades and storage, particularly in areas with high solar irradiation in the south-central coast of Vietnam (Dapice 2018). The government should also consider implementing a carbon price for certain energy-intensive industries such as cement and steel production.

If no course correction is made to the current Power Development Plan adopted by the Vietnamese government in 2016, up to 20,000 deaths per year are expected by 2030 due to exposure to particulate matter pollution from coal-fired power plants (Koplitz et al. 2017). Another concern is that Vietnam has started importing coal since 2013, exposing itself to energy security risk due to dependence on coal supply and imported equipment. Coal-fired plants also emit the largest amount of greenhouse gases and contribute to climate change, and the Mekong River Delta of Vietnam is one of the regions most vulnerable to sea level rise, tropical cyclones, and seasonal river flows. Vietnam's energy sector is also affected by actions taken by other countries in the region, particularly with regard to damming the Mekong River for hydropower generation by Lao PDR and China in the Upper Mekong Basin. Climate change may alter water flow patterns, potentially affecting hydropower generation, which the Vietnamese government is increasingly looking at as a cheap source for electricity

imports. In foreseeing these threats and addressing excessive energy dependence appropriately, the economic, environmental, and security benefits of a less energy-intensive development trajectory will be significant.

Supplementary information

Supplementary information accompanies this paper at- https://doi.org/10.1186/s40008-019-0168-9.

Additional file 1. Additional tables.

References

ADB: Vietnam energy sector assessment, strategy, and road map(2016) Asian Development Bank, Metro Manila, Philippines

Australian Energy Council: electricity prices around the world: what is the impact of renewable charges? (2016) Australian Energy Council

Berndt E, Wood D (1975) Technology, prices, and the derived demand for energy. Rev Econ Stat 3:259–68

Burke PJ, Csereklyei Z (2016) Understanding the energy-gdp elasticity: a sectoral approach. Energy Econ 58:199–210

Chung JW (1987) On the estimation of factor substitution in the translog model. Rev Econ Stat 69(3):409–17

Dapice D (2018) Vietnam's crisis of success in electricity: Options for a successful clean energy mix. Kennedy School of Governance, Harvard University, Boston, MA, Ash Center for Democratic Governance and Innovation

Dapice D, Le PV (2018) Counting all of the costs: Choosing the right mix of electricity sources in vietnam to 2025. In: Agriculture, Livelihoods, and the Environment in the Lower Mekong Basin. SIRD, Malaysia

Dargay JM (1983) The demand for energy in swedish manufacturing industries. Scand J Econ 85(1):37–51

FAO: Cost control in forest harvesting and road construction (1992)Food and Agriculture Organization of the United Nations, Rome, Italy

FPT: Electricity sector report: a competitive market message (2015) FPT Securities, Hanoi, Vietnam

Garg V, Bridle R, Clarke K (2015) Energy pricing, energy supply and fdi competitiveness in viet nam: an assessment of foreign investor sentiment. The International Institute for Sustainable Development, Viet nam

GIZ: Opportunities to enhance energy efficiency in vietnamese firms (vietnamese) (2017) Gesellschaft fur Internationale Zusammenarbeit, Hanoi, Vietnam

Greer M (2012) Chapter 9: Price and substitution elasticities of demand: how are they used and what do they measure? In: Greer M (ed) Electricity marginal cost pricing, applications in eliciting demand responses. Butterworth-Heinemann, Oxford

Griffin JM, Gregory PR (1976) An intercountry translog model of energy substitution responses. Am Econ Rev 66(5):845–57

Hartman RS (1979) Frontiers in energy demand modeling. Ann Rev Energy 4:433–66

Kleijweg A, Huigen R, van Leeuwen G, Zeelenberg K (1990) Firm size and the demand for energy in dutch manufacturing, 1978–1986. Small Bus Econ 2:171–81

Koplitz SN, Jacob DJ, Sulprizio MP, Myllyvirta L, Reid C (2017) Burden of disease from rising coal-fired power plant emissions in southeast asia. Environ Sci Technol 51:1467–76

Ministry of Investment and Trade: Energy Outlook Report 2017. Hanoi, Vietnam (2017) Ministry of Investment and Trade

Pindyck RS (1979) Interfuel substitution and the industrial demand for energy: an international comparison. Rev Econ Stat 61(2):169–79

Pitt MM (1985) Estimating industrial energy demand with firm-level data: the case of indonesia. Energy J 6(2):25–39

Statistica: primary energy consumption in Vietnam from 2005 to 2017 (2019) Statistica

Tang CF, Tan BW, Ozturk I (2016) Energy consumption and economic growth in Vietnam. Renew Sustain Energy Rev 54:1506–14

Thompson H (1997) Substitution elasticities with many inputs. Appl Math Lett 10(3):123–27

Thompson H (2006) The applied theory of energy substitution in production. Energy Econ 28:410–25
United Nations Development Programme: Value Chain and Policy Analysis of Fossil Fuel Trade, Subsidy and Tax in Viet
        Nam. Hanoi, Vietnam (2011) United Nations Development Programme
United Nations Development Programme: Green Growth and Fossil Fuel Fiscal Policies in Viet Nam - Recommendations
        on a Roadmap for Policy Reform. Hanoi, Vietnam (2014) United Nations Development Programme
Uri ND (1979) Interregional energy substitution in manufacturing in the united states. Energy Conver 19:25–31
Woodland AD (1993) A micro-econometric analysis of the industrial demand for energy in nsw. Energy J 14(2):57–89
World Bank (2019) Energy intensity level of primary energy. World Bank, Geneva

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.