Productivity dispersion and the roles of quality of labour input and competition: A case of Vietnamese manufacturing sector

Doan, Tinh and Nguyen, Ha

Ministry of Business, Innovation and Employment, New Zealand, Victoria University of Wellington, New Zealand

16 July 2013

Productivity dispersion and the roles of quality of labour input and competition: A case of Vietnamese manufacturing sector

Tinh Doan*

Ministry of Business, Innovation and Employment, New Zealand

Ha Nguyen

School of Economics and Finance
Victoria University of Wellington, New Zealand

Abstract

This paper examines the roles of cost of labour input and competition on productivity dispersion in the Vietnamese manufacturing sector. We look at the effect accounting for labour input quality has on explaining productivity dispersion. This paper tests the hypothesis that mismeasurement of labour input may play a role in large productivity dispersion. We use the cost of labour input of firms as a proxy measure of labour input quality to examine whether incorporating this measure accounts for a part of the productivity dispersion. The paper also examines the role of competition in the extent of productivity dispersion.

JEL Classification: L11, L25, J24, P27

Keywords: *productivity dispersion, competition, labour input, transition economies*

---

* Correspondence: Tinh Doan, Ministry of Business, Innovation and Employment, New Zealand; e-mail: thanhtinhdoan@gmail.com

1. Introduction

There is a rich literature on productivity dispersion elsewhere over the world, but there is no evidence on this topic is observed in Vietnam. This motivates the current paper to examine this issue in the Vietnamese manufacturing sector. Our paper has two main focuses. First, we compute the productivity dispersion using labour and total factor productivity, then consider whether using a quality measure of labour input can explain a part of productivity dispersion within the Vietnamese manufacturing sector. This is because mismeasurement of inputs due to unavailability of data may increase the measurement errors across the distribution, and hence inflate the dispersion. Second, we propose some possible explanations to productivity dispersion including the role of competition. To do so in this paper we estimate production functions for separate industries using both measure of labour count input and cost of labour input. The two different measures are used to determine how important the mismeasurement of inputs and if this increases productivity dispersion. We find that using cost of labour input in the production function considerably reduces the dispersion. We then suggest using cost of labour input to estimate production function in the case of the Vietnamese manufacturing sector.

The remainder of this paper is structured as follows. The next section reviews the previous work to provide a background for this paper. In section three we outline the method and data used in the analysis. Section four presents the results. Section five provides some international comparisons and proposes some potential explanations to productivity dispersion in the Vietnamese manufacturing sector. Final section is for summary and conclusions.

2. Productivity dispersion and the role of labour input quality

Large productivity dispersion are common across industries and firms (Syverson, 2011; Bartelsman and Doms, 2000). A commonly asked question is that why the large dispersion even within very narrowly defined industries exists persistently when the competition theory predicts poor performers in a market to become unprofitable and eventually be selected out of the market. New entrants are expected to be more productive or at least more productive than a certain group of firms in the market such as exiters as they come into the market with a new idea or innovation displacing the low productivity incumbents. This selection process is expected to increase the overall productivity of the market and reduce the dispersion over time. Syverson (2004b) states that productivity dispersion decreases in more competitive markets, where the selection process is evident.

Despite these selection mechanisms, productivity dispersion remains large and persistent across industries and many countries including well-developed market countries. Much of the literature on this field tries to explain the large differences in productivity. Firms are typically heterogeneous, own different inputs (Griliches, 1957) including firm characteristics such as management or business practices or technology. Competition may not be enough to sort out poor performers. Understanding how the poor performance firms differ from the top performers influence how we think about the allocation of resources to the most productive use within industry or across industries.

Given the same labour input count, more productive firms tend to pay higher wage rates (Bagger *et al*, 2010), and in a well function labour market only better quality (education, skilled, ability, motivation) workers can earn higher wage, and in turn higher paid workers will work harder. This is called ‘rent sharing’ (Bagger *et al*, 2011). Better quality workers work for higher paying firms. This suggests that cost of labour input can better capture labour input quality and a better measure of labour input.

Therefore, incorporating the cost of labour input as a proxy for labour input quality may reduce the productivity dispersion. It means that part of the productivity dispersion using labour input count reported is an artefact of mismeasurement of labour input, or productivity dispersion is overestimated (Bagger, Christensen and Mortensen, 2011; Fox and Smeets, 2011). In this paper we examine labour input quality to determine if it makes up some of the productivity dispersion.

The quality of labour input proxied by cost of labour input or wage bill is better used in the production function for at least two reasons. First, capital input is typically measured in monetary units, whereas labour inputs are often measured by volume such as number of workers or hours worked. Labour input or human capital is difficult to quantify as they rely on not only measures of each individual worker’s skill, experience and education but also unobservable attributes like ability and motivation. Cost of labour input can capture all this factors under a well-functioned labour market.

The wage bill can represent labour quality as wage rate is price of labour or indicator of labour quality. The wage bill well captures the unobservable characteristics of the worker and firm. For example, Hyslop and Mare (2006) in a study of New Zealand firms find about half of worker's income variation is due to worker fixed effects, 10 to 25 percent is due to firm fixed effect, while only a quarter is attributed to workers' observed characteristics such as age and gender.

Second, cost of labour input may better reflect labour input in the case of the Vietnamese manufacturing sector because the Vietnam Enterprises Survey (VEC) which provides us data used in this paper does not provide sufficient information on working status: full-time and part-time labour count, under-time and over-time hours worked, while these labour market situations are common in the country. For example, the state owned enterprises (SOEs) have retained full-time workers but their staffs often underperform,[2] have low productivity and then earn low wage, while private sectors often ask their staffs/workers to work over-time and on weekends to maximize hours of machinery operations in order to minimize cost (depreciation unit cost) and to meet their timeline of contract delivery. In such circumstances, labour count fails to capture contribution of labour input.

In the recent decades, the increasing availability of firm-level datasets has led to an increasing number of studies exploring the productivity differences across firms. The key mains of this sort of research is to analyse difference in inputs and outputs across firms (Fox and Smeets, 2011; Bagger, Christensen and Mortensen, 2011; Ornaghi, 2006; Foster, Haltiwanger and Syverson, 2008), and to determine factors affecting the productivity dispersion. Most common considered factors are competition, exit and entry of firms, trade, and regulations, technology advancement, development of ICT, and productivity spillovers. For example, Syverson (2004a) looks at relationship between competition and productivity dispersion, Foster Haltiwanger and Krizan, (2006) look at effects of entry and exit on productivity; Helpman (2006), Melitz (2003), and Helpman, Melitz and Yeaple (2004) look at effects of trade on the dispersion; Nicoletti and Scarpetta (2005) consider regulation and institutions, while Faggio, Salvanes and Van Reenen (2010) examine

[2] As efficiency is not their first target of businesses, the firms often claim their losses due to pursuing political targeting. Therefore, SOEs still receive subsidies from the government under some forms, and firing workers are somehow restricted by these business ideology.

technology improvements, and Bloom, Schankerman and Van Reenen (2007) study productivity spillovers.

There is some international evidence on the relationship between labour quality and productivity dispersion. Hellerstein and Neumark (2006) look at wage in the production function and report a reduction in the dispersion through incorporation of the wage as a labour quality measure. Fox and Smeets (2011) use an employer-employee linked dataset to create human capital variables that capture education and experience of employees. Their findings indicate the cost of labour input explains as much productivity dispersion as the human capital measures. A similar study by Irarrazabal, Moxnes and Ulltveit-Moe (2009) uses similar human capital measures as Fox and Smeets (2011) to look at the productivity of internationalising firms and find that inclusion of labour input quality explains 25 precent of productivity differences between internationalises and non-internationalises which overstates impact of internationalization on the productivity differences. An alternative method used by Bagger *et al* (2011) incorporates occupational data with the cost of labour input to augment the production function. Productivity dispersion across industries reduces but not much. These studies are for developed countries, very limited evidence is observed so far for developing economies.

3. Measuring productivity and data sources

In this paper total factor productivity (TFP) are obtained from Cobb-Douglas using OLS and FE estimators. The conventional measure of labour input and the quality labour input measure, proxied by the cost of labour input will be used with either of the production function estimators.

Cobb-Douglas production function is as follows:

$$Lny_{it} = \beta_0 + \beta_l \, lnl_{it} + \beta_K lnk_{it} + \beta_{ic} lnic_{it} + v_{it} + e_{it} \qquad (1)$$

where y_t is gross output (sales), l is the number of workers including paid working proprietors, k is capital or cost of capital services, ic is intermediate consumption such as material, power bill, water bill, and β_l, β_k and β_{ic} are the estimated coefficients on labour input variable, capital and intermediate consumption, respectively. The residual ($v_{it} + e_{it}$) is measured as TFP. The residual captures productivity that is not explained by either measured inputs. Although the notation subscript j (industry) has been left out for simplicity this production function for each firm i in each two-digit industry j at time t.

We then replace the labour count (l) with the cost of labour input (w). Our dataset offers both labour count and cost of labour input for the firms. The cost of labour input is calculated as the sum of a firm's total wages and salaries including other employees' incomes such as holiday allowance, capturing both full time and part time employees including paid working proprietors (i) in year t.

Some of inputs such as managerial ability, R&D stocks, intangible assets, quality of inputs (land, labour, capital, business location), expected economic downturn, expected weather etc. are unobserved (v_{it}), in fact, unmeasured by econometricians, but are observed (or predictable) by firms. These factors are called unobserved productivity shocks. Therefore, firm's optimal choice of inputs l_{it} and k_{it} will be correlated with v_{it}. This is a correlation between unobserved productivity shocks and input usage levels. Firms will increase output to maximize profits if they face positive productivity shocks, so they will increase the uses of inputs. Thus, the observed/measured inputs are endogenous, in other words there is simultaneity problem here causing OLS estimates of β's to be biased (Ackerberg, Caves & Frazer, 2006).

What are solutions to this issue? First is instrument variable approach. This method needs valid instruments that are correlated with endogenous variable (input level choice, e.g. labour) but not correlated with firm outcome or its residual (error terms). It is typically hard to find good instruments that satisfy the conditions. Input prices (interest rate and/or wage rate) can be potential instruments, but input prices are often unavailable in datasets or do not vary or do not vary enough across firms. Even if there is a variation in input price, it may accounts for market power in input markets or heterogeneity in quality of inputs e.g. worker quality, that may invalidate the use of input price as an instrument (Ackerberg, Caves & Frazer, 2006). Recently, some suggest using Levinsohn and Petrin (2003), an extension of Olley and Pakes (1996), this approach uses firm's intermediate consumption to control for the endogeneity. However, there are some disadvantages of these approaches such as identification and estimation issues (see more detailed discussion in Ackerberg et al, 2006 and Wooldridge, 2009). The lack of exogenous instruments to address the endogeneity and lack of outperformed methods, we therefore will apply conventional approach OLS and FE to Cobb-Douglas production function in this paper to estimate TFP. Fixed effect estimator assumes that unobserved productivity shocks (v_{it}) for each firm is time-invariant. That is $v_{it} = v_{i,t-1}$ for all t. Instead of estimating original values as does the OLS, the FE estimates deviations from individual means. Since mean of v_{it} is itself so deviation from its mean will be zero, then is removed.

The data used in this paper comes from the Vietnam Enterprise Census (VEC) conducted annually by General Statistics Office Vietnam since 2000. The VEC offers a panel dataset spanning from 2000 to 2009. All registered firms have to fill the questionnaire provided by district statistics offices as legal liability described in the Vietnam Statistical Law.[3] The VEC provides comprehensive information about firms and their activities in the first decade of the 21st century. The census offers information on firm demographics, business activities, employment, wages, assets, capital, business performance, cost, revenue, and profit.[4]

Industries have been defined in this paper by the Vietnam Standard Industrial Classification 1993 (VSIC1993) two-digit industry level codes. Production function estimation has been done at this level. This may be the narrowest level to estimate productivity as we move to lower level we may lose many industries because of insufficient observations.

Our analysis is restricted to the manufacturing sector that may offer the best quality data amongst all industries in Vietnam as well as elsewhere. We removed firms without tax code for some reasons such as missing data or infant firms without tax codes since we use the tax code as firm identifiers to merge data. We also dropped extreme observations in the bottom one percentile and the top one percentile. This may reduce the productivity dispersion, but makes our data less noisy.

4. Results

In this section, we first present the results from estimating productivity dispersion within two-digit industries. Labour productivity distribution, the most commonly used, is presented first to establish the initial picture of productivity dispersion. We then

[3] The GSO has offices in all districts and provinces. The district offices report directly to provincial statistics offices.
[4] More information about the variables used in this paper, see Appendix 6.

estimate TFP using two estimators OLS and FE to provide the extent to which the two alternative uses of firm labour inputs to explain productivity dispersion. Two measures, ratio of 90th percentile over 10th percentile (P90/10) and standard deviation, of productivity distribution will be used to test the role of labour input quality in explaining the extent of dispersion.

4.1 Labour productivity distribution

Table 1 shows the labour productivity (LP) distributions presented at the two-digit VSIC 1993. The 90/10 ratios show the difference between firm labour productivity at the top 90th percentile and firms at the bottom 10th percentile of productivity distribution. [5] We also presented P90/50 and P50/10 ratios to provide more information about the dispersion on the left and right halves of the distribution. This helps indicate the extent of any skew towards one end or the other on the distribution.

Table 1 shows that there is a large variation of productivity within industries and cross industries. The 'Coal, petroleum and biofuel product production', 'Radio, TV and telecommunication equipment manufacturing', 'Chemical and chemical product manufacturing', 'Office and computer manufacturing' and 'Vehicle and trailer manufacturing' industries are highly dispersed (columns 1 and 2). The firms at the 90th percentile can be up to 35 times as productive as firms at the 10th percentile in 'Coal, petroleum and biofuel product production'. The dispersion is much lower but still high (5 times or more) in the 'Clothing, leather dying and manufacturing', 'Leather tanning and fur dressing, luggage & other leather product manufacturing', 'Paper and paper products', and 'Wood processing and wooden product manufacturing'. It seems that sectors those are labour intensive are less dispersed than the capital intensive sectors. This is also found in many countries even in market mature economies such as New Zealand (Devine *et al*, 2012). The productivity distribution is even higher when we include extreme outliers and firms with less than one employee.[6] On average, the P90/10 is 8.3-to-1, much higher than observed countries such as the US, UK, Denmark (Syverson, 2004b; Haskel and Martin, 2002; Fox and Smeets, 2011).[7]

[5] The dispersion is first estimated at four-digit level for each year then aggregated up to two-digit level

[6] These may be part-year or part-time single-staff firms.

[7] No similar information about developing or transition economies is available to compare, this may make the comparison less sensible

The P90/50 and P50/10 distributions do not vary in a great deal. There is a greater variation in the upper half of the distribution (P50/10 column) compared with the lower half of the distribution, suggesting a wider spread amongst higher productivity performers. Fifteen out of the 23 industries have larger ratios (P90/50) than P50/10. The overall economy has a greater dispersion of productivity in the upper half of the distribution, 2.84-to-1 versus 2.8-to-1 in the lower half of the distribution. This suggests that the distribution is slightly productivity-skewed to the right, but overall the productivity is quite symmetrically distributed.

The standard deviation (SD) provides more information about highly dispersion. Standard deviation shows how much variation or dispersion exists from each industry's mean. A high standard deviation indicates that the data points spread out over a large range of values. The SD columns show that sectors that have higher means of P90/10 ratios typically have higher standard deviations. But for some industries such as the 'Printing, publishing, and recording', and 'Medical & precious tools & equipment' despite lower means the standard deviations are extremely high because some component (four-digit level) industries within two-digit industries the (P90/10) ratios do spread out massively.[8]

Given the large dispersion of labour productivity, a question arises here is that what drives the large dispersion, whether cost of labour input or wage rates play any role in such high dispersion? We now turn out to look at Total Factor Productivity (TFP) to answer the question. To do so we compute TFP using two alternative measures of labour input: labour count and cost of labour input or wage bill.

TFP can be estimated using different specifications and models. Amongst others we use the conventional Cobb-Douglas function OLS and Fixed Effect (FE) estimator. These enable us to account for effects of both capital and labour inputs. The measure of labour input quality can be captured by using the cost of labour input in monetary terms as discussed earlier. We then compare changes in TFP distribution P90/10 ratios (both means and standard deviations).

[8] Bear in mind that the ratios are first computed at 4-digit industry level for each year.

Table 1: Labour Productivity Distribution of manufacturing sector, 2000-2009

Two-digit industry (VSIC1993)	P90/10		P50/10		P90/50	
	Mean	SD	Mean	SD	Mean	SD
Food and drink manufacturing	8.85	8.54	2.84	0.96	2.97	1.34
Tobacco and cigarettes manufacturing	8.71	1.78	2.36	0.37	3.71	0.76
Textile manufacturing	10.75	4.51	3.70	1.07	2.92	0.74
Clothing, leather dying and manufacturing	5.10	2.10	2.60	0.33	1.94	0.63
Leather tanning and fur dressing, luggage & other leather product manufacturing	5.98	4.59	2.58	1.04	2.25	0.57
Wood processing and wooden product mfg	6.53	1.31	2.87	0.44	2.28	0.30
Paper and paper products	5.92	1.29	2.33	0.28	2.53	0.38
Printing, publishing, and recording	6.42	19.10	2.51	4.49	2.28	1.18
Coal, petroleum product & biofuel product mfg	34.58	37.70	3.68	3.40	9.94	7.97
Chemical and chemical product manufacturing	17.38	11.47	3.60	1.68	4.68	1.70
Rubber and plastic product production	7.98	5.67	2.73	1.38	2.89	0.64
Glass, glass product & fine ceramic product mfg	8.00	4.35	2.70	0.63	2.87	1.08
Metal manufacturing	11.50	8.15	2.62	0.84	4.36	2.64
Metal product manufacturing	6.62	2.35	2.51	0.55	2.61	0.64
Machinery and equipment manufacturing	7.76	7.98	2.82	2.45	2.70	1.20
Office and computer equipment manufacturing	15.43	14.44	3.23	1.62	4.68	2.99
Electrical and machinery equipment mfg	11.81	8.11	3.03	1.17	3.88	2.10
Radio, TV and telecoms equipment mfg	20.16	14.50	3.46	1.79	5.89	3.33
Medical & precious tools & equipment	10.38	26.88	3.59	8.93	3.00	1.16
Vehicle and trailer manufacturing	13.34	13.82	3.09	1.82	4.23	4.30
Other transport equipment manufacturing	10.81	7.20	2.99	0.81	3.51	1.65
Bed, wardrobe, table, chair & other wooden product manufacturing	6.21	3.70	2.62	0.88	2.32	0.69
Recycling	6.90	3.40	2.76	1.55	2.52	0.75
Overall manufacturing	**8.30**	**8.37**	**2.80**	**1.56**	**2.84**	**1.44**

Note: Labour productivity is calculated as ln(VA) - ln(labour). Values are calculated for the period 2000-2009. Top one and bottom one percentiles were removed, and firms with less than one employee were removed.

4.2 Cost of labour input and reduction of TFP dispersion

Table 2 reports the results of the FE estimates of Cobb-Douglas production function, estimated for 23 two-digit VSIC 1993 industries. [9] Results are presented using both the conventional measure for labour input (labour count), and the cost of labour input for each two-digit industry. Estimations have been done for each two-digit industry to allow for different technologies.

Overall the regression results presented in Table 2 are what we expect with positive and significant coefficients for all capital (K), intermediate consumption (IC) and

[9] The OLS results generally follow a similar pattern, see Appendix 3.

labour input (L) measures. Bear in mind that this paper does not come to any inference on the production function estimators and whether better specification of inputs provides better specification of productivity measures. We assume that simultaneity bias if it exists is in the same direction for both models with labour count and with cost of labour input.

For each industry, model with labour count and cost of labour input are presented in left and right column respectively. Almost all coefficients are positive and highly significant at the one per cent level, but as we run FE model, some industries such as industries 17 and 30 have few observations, just above 20 observations per year, hence the coefficients are not statistically significant. The coefficients on labour count input in row one are relatively higher in labour intensive industries such as 'Leather tanning and fur dressing, luggage and other leather product manufacturing' and 'Clothing, leather dying and manufacturing' while the coefficient of capital is highest in capital-intensive industries such as industry 26 'Glass, glass product and fine ceramic product production'. Meanwhile the lowest coefficient of IC is found in 'Leather tanning and fur dressing, luggage and other leather product manufacturing' and highest in 'Coal, petroleum product and biofuel product production' (Table 2).

When the cost of labour input is included in the production function in place of the labour count the coefficients of labour input generally increase. This implies that cost of labour input better captures contribution of labour in firm's total output. The (within) R-squared has slightly improved when using wage as labour input.

Table 2: Cobb-Douglas Production Function (FE estimator), 2000-2009

Dependent variable Log VA	D15		D16		D17		D18		D19		D20		D21		D22	
	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage
lnL	0.148**	0.178**	0.117	0.095*	0.250**	0.276**	0.356**	0.402**	0.388**	0.430**	0.185**	0.229**	0.162**	0.191**	0.246**	0.245**
	[0.006]	[0.005]	[0.076]	[0.041]	[0.015]	[0.013]	[0.019]	[0.014]	[0.033]	[0.024]	[0.010]	[0.008]	[0.017]	[0.016]	[0.017]	[0.014]
lnK	0.138**	0.123**	-0.007	-0.011	0.141**	0.125**	0.116**	0.091**	0.140**	0.115**	0.100**	0.086**	0.103**	0.093**	0.093**	0.088**
	[0.004]	[0.004]	[0.040]	[0.039]	[0.012]	[0.011]	[0.009]	[0.007]	[0.016]	[0.015]	[0.007]	[0.006]	[0.009]	[0.009]	[0.007]	[0.007]
lnic	0.683**	0.665**	0.711**	0.699**	0.590**	0.564**	0.504**	0.473**	0.479**	0.463**	0.662**	0.622**	0.702**	0.673**	0.661**	0.638**
	[0.006]	[0.006]	[0.099]	[0.102]	[0.014]	[0.014]	[0.011]	[0.011]	[0.022]	[0.020]	[0.011]	[0.011]	[0.021]	[0.021]	[0.017]	[0.016]
Constant	1.585**	1.349**	2.798**	2.810**	1.807**	1.444**	2.078**	1.389**	2.007**	1.164**	1.627**	1.418**	1.483**	1.269**	1.539**	1.162**
	[0.034]	[0.031]	[0.886]	[0.880]	[0.094]	[0.095]	[0.074]	[0.075]	[0.165]	[0.168]	[0.051]	[0.045]	[0.106]	[0.095]	[0.081]	[0.071]
Observations	41411	41411	233	233	7992	7992	13524	13524	4263	4263	14137	14137	7997	7997	9803	9803
R-squared	0.90	0.91	0.92	0.93	0.86	0.88	0.82	0.86	0.80	0.85	0.90	0.91	0.92	0.93	0.92	0.93

Dependent variable Log VA	D23		D24		D25		D26		D27		D28		D29		D30	
	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage
lnL	0.083*	0.085*	0.179**	0.187**	0.213**	0.210**	0.263**	0.296**	0.111**	0.112**	0.193**	0.211**	0.266**	0.259**	0.512**	0.12
	[0.038]	[0.037]	[0.020]	[0.014]	[0.017]	[0.014]	[0.015]	[0.011]	[0.016]	[0.015]	[0.010]	[0.009]	[0.039]	[0.027]	[0.185]	[0.172]
lnK	0.117**	0.104**	0.088**	0.082**	0.103**	0.097**	0.157**	0.139**	0.084**	0.080**	0.097**	0.090**	0.108**	0.105**	0.133	0.247
	[0.034]	[0.030]	[0.007]	[0.007]	[0.009]	[0.009]	[0.007]	[0.007]	[0.013]	[0.013]	[0.005]	[0.005]	[0.013]	[0.012]	[0.116]	[0.149]
Lnic	0.838**	0.818**	0.732**	0.715**	0.699**	0.679**	0.550**	0.516**	0.792**	0.774**	0.700**	0.673**	0.665**	0.639**	0.665**	0.779**
	[0.032]	[0.036]	[0.017]	[0.017]	[0.015]	[0.015]	[0.013]	[0.013]	[0.021]	[0.023]	[0.010]	[0.010]	[0.026]	[0.027]	[0.097]	[0.150]
Constant	0.581+	0.615+	1.306**	0.992**	1.342**	1.052**	1.879**	1.517**	1.158**	1.085**	1.421**	1.124**	1.395**	1.036**	0.552	-0.22
	[0.330]	[0.321]	[0.115]	[0.107]	[0.089]	[0.088]	[0.070]	[0.064]	[0.097]	[0.091]	[0.049]	[0.044]	[0.141]	[0.131]	[0.581]	[0.810]
Observations	166	166	8505	8505	11647	11647	15067	15067	3549	3549	20360	20360	5441	5441	211	211
R-squared	0.97	0.97	0.91	0.92	0.91	0.91	0.86	0.88	0.95	0.95	0.92	0.93	0.87	0.88	0.89	0.87

Dependent variable Log VA	D31		D32		D33		D34		D35		D36		D37	
	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage
lnL	0.196**	0.196**	0.254**	0.262**	0.585**	0.582**	0.272**	0.238**	0.260**	0.271**	0.247**	0.285**	0.077**	0.122**
	[0.030]	[0.029]	[0.036]	[0.025]	[0.203]	[0.153]	[0.042]	[0.031]	[0.031]	[0.020]	[0.013]	[0.010]	[0.029]	[0.030]
lnK	0.116**	0.110**	0.096**	0.083**	0.05	0.02	0.123**	0.113**	0.127**	0.107**	0.100**	0.088**	0.124**	0.112**
	[0.015]	[0.013]	[0.020]	[0.020]	[0.057]	[0.052]	[0.017]	[0.018]	[0.014]	[0.012]	[0.007]	[0.006]	[0.030]	[0.027]
Lnic	0.721**	0.703**	0.727**	0.716**	0.684**	0.641**	0.674**	0.660**	0.637**	0.612**	0.639**	0.601**	0.752**	0.731**
	[0.029]	[0.029]	[0.028]	[0.027]	[0.080]	[0.075]	[0.029]	[0.029]	[0.022]	[0.022]	[0.012]	[0.012]	[0.039]	[0.039]
Constant	1.201**	0.914**	1.083**	0.553**	0.567	-0.452	1.243**	0.993**	1.582**	1.222**	1.551**	1.157**	1.091**	0.902**
	[0.176]	[0.188]	[0.201]	[0.194]	[0.571]	[0.680]	[0.193]	[0.177]	[0.154]	[0.141]	[0.065]	[0.059]	[0.155]	[0.128]
Observations	3184	3184	1751	1751	787	787	2545	2545	4314	4314	13740	13740	371	371
R-squared	0.90	0.89	0.88	0.81	0.78	0.90	0.89	0.88	0.86	0.90	0.88	0.96	0.96	0.96

*Notes: Robust standard errors in brackets; + significant at 10%; * significant at 5%; ** significant at 1%; All models controlled for year fixed effect. lnL is natural log of labour count or natural log of Wage depending on models with labour count or with cost of labour input. lnK is natural log of cost of capital. To have comparable results, both labour count and wage model use the same sample where both missing lnL and missing lnwage are dropped. For industry description, see Appendix 1.*

We now use the P90/10 ratio to examine whether including the labour input quality (proxied by cost of labour input) explains some of the dispersion in productivity. The dispersion is calculated using percentile differences in TFP which is obtained from the residuals of Cobb-Douglas production functions.[10] Table 3 provides dispersion estimates using Cobb-Douglas production function. Again we present only Cobb-Douglas FE results, results from the OLS models are reported in the Appendix 2 as the results follows a very similar pattern. Columns 2-3 report results using labour count and columns 4-5 outline the results using the cost of labour input. The last two columns present the differences between these two specifications for the means and standard deviations.

Broadly we find some key features when we compare the changes in dispersion across industries. TFP dispersion is much lower than LP dispersion noted previously. The 90th percentile firms' workers can be up to 35 times as productive as the 10th percentile firms' workers in 'Coal, petroleum and biofuel product production', even the lowest LP dispersion is around 6-to-1, which is much higher than the TFP dispersion at 1.96-to-1 and 1.78-to-1 for model with L and wage respectively. The role of capital in production function estimation may help explain a sharp decline in the TFP dispersion why LP fails to do so. Interestingly, we observe relatively high dispersion in high technological industries of 'Office and computer manufacturing' and 'Medical, precious tools & equipment' where intellectual properties may hinder market entry, then competition and preserve larger productivity gaps between firms with different innovation levels.

Overall we find that including a quality measure of labour input, the cost of labour input, reduces the observed productivity dispersion by 0.15 points or 8.7%. This pattern is observed in almost all industries except 'Tobacco and cigarettes manufacturing'. There is not a very large dispersion across firms suggesting that competition in the economy is relatively high and there is an easy mobility of resources across sectors.

There are remarkable decreases in P90/10 ratios for 'Leather tanning and fur dressing, luggage and other leather product manufacturing', 'Clothing, leather dying and manufacturing', 'Printing, publishing, and recording', 'Chemical and chemical

[10] The FE residual is the combined residual consisting of the fixed-error and overall error components

product production', 'Glass, glass product and fine ceramic product production', 'Office and computer equipment production', 'Electrical and machinery equipment production', 'Radio, TV and telecommunications equipment production', 'Medical & precious tools & equipment', and 'Bed, wardrobe, table, chair and other wooden product manufacturing'. The decline rates are more than nine per cent. We observe that these sectors tend to be non-state sectors or high tech and pay their workers based on individual productivity or performance. Therefore workers' income well capture workers' experience, skills, effort, hours worked and work motivation rather than workers' presence and their education which is widely believed low quality in Vietnam. This may imply that the labour input if mismeasured can result in biases in the calculated dispersion especially for more labour market-oriented industries. The cost of labour input appears to reflect partly price of labour by which workers are paid their marginal value. The pattern of change in the dispersion due to cost of labour input used does hold for the OLS model (see Appendix 2).

Fox and Smeets (2011) and Bagger *et al*. (2011) who indicate that using quality measures of labour input reduces the dispersion but not much in the study developed economies. Our finding of reduced dispersion by using cost of labour input is consistent with the emerging literature, but a significant decline in dispersion when using cost of labour input in our case is contradicting the literature on the developed countries.

Table 3: Total Factor Productivity Distribution (FE model of Cobb-Douglas with labour count input vs. cost of labour input)

Two-digit industry (VSIC1993)	L		Wage		Difference		
	P90/10		P90/10		P90/10		
	mean	SD	Mean	SD	Mean	SD	%
Food and drink manufacturing	2.20	3.13	2.09	2.91	-0.11	-0.22	-5.0
Tobacco and cigarettes manufacturing	2.86	0.30	3.04	0.37	0.18	0.07	6.3
Textile manufacturing	2.12	0.35	1.98	0.26	-0.13	-0.09	-6.3
Clothing, leather dying and manufacturing	2.21	0.67	1.86	0.83	-0.36	0.16	-16.1
Leather tanning and fur dressing, luggage and other leather product manufacturing	2.51	0.41	2.09	0.59	-0.42	0.18	-16.8
Wood processing and wooden product mfg	1.91	0.11	1.77	0.12	-0.15	0.01	-7.6
Paper and paper products	1.60	0.12	1.50	0.08	-0.09	-0.03	-5.9
Printing, publishing, and recording	1.55	0.30	1.41	0.23	-0.14	-0.07	-9.1
Coal, petroleum product and biofuel product production	1.58	0.55	1.52	0.49	-0.06	-0.06	-4.1
Chemical and chemical product production	1.77	0.19	1.58	0.16	-0.18	-0.03	-10.3
Rubber and plastic product production	1.63	0.14	1.55	0.13	-0.08	-0.01	-4.9
Glass, glass product and fine ceramic product production	2.15	0.29	1.80	0.21	-0.35	-0.08	-16.3
Metal manufacturing	1.53	0.12	1.53	0.14	0.00	0.03	-0.2
Metal product manufacturing	1.67	0.18	1.56	0.15	-0.11	-0.03	-6.6
Machinery and equipment production	1.73	0.42	1.62	0.62	-0.11	0.19	-6.1
Office and computer equipment production	6.42	3.43	5.59	8.06	-0.83	4.63	-12.9
Electrical and machinery equipment production	1.77	0.29	1.60	0.26	-0.17	-0.04	-9.4
Radio, TV and telecommunications equipment production	2.05	0.48	1.83	0.39	-0.22	-0.09	-10.7
Medical & precious tools & equipment	4.15	5.33	3.60	7.14	-0.55	1.81	-13.3
Vehicle and trailer manufacturing	1.80	0.24	1.65	0.29	-0.15	0.05	-8.1
Other transport equipment manufacturing	2.01	0.29	1.84	0.28	-0.17	-0.01	-8.6
Bed, wardrobe, table, chair and other wooden product manufacturing	1.91	3.17	1.69	2.24	-0.22	-0.93	-11.3
Recycling	1.49	0.26	1.38	0.19	-0.11	-0.06	-7.3
Overall manufacturing	**1.96**	**1.77**	**1.78**	**1.63**	**-0.17**	**-0.15**	**-8.7**

Notes: Top one and bottom one percentiles at 4-digit level are removed before computing these numbers. The production function estimation controlled for time fixed effect.

Now we look at the whole TFP distribution (Figure 1) and its standard deviations (Table 4) using both measures of labour input to see whether using the cost of labour input (wage) narrows down the TFP dispersion. The TFP dispersion and its standard deviation (SD) are first computed at four-digit industry level, we report here results for aggregate levels of two-digit industry SD.[11] Overall the results are consistent across methods of production function estimation, and for the conciseness reason we report results from the FE models of the T-test for standard

[11] The standard deviation is a measure of how spread out numbers from the mean of TFP. Standard deviation is the square root of the variance. The variance is defined as the average of the squared differences from the mean.

deviation differences which we believe more precise and conclusive. The similar results from OLS allow testing for more current years (2008-2009) which may eliminate the dispersion caused by time effect are reported in Appendix 4.

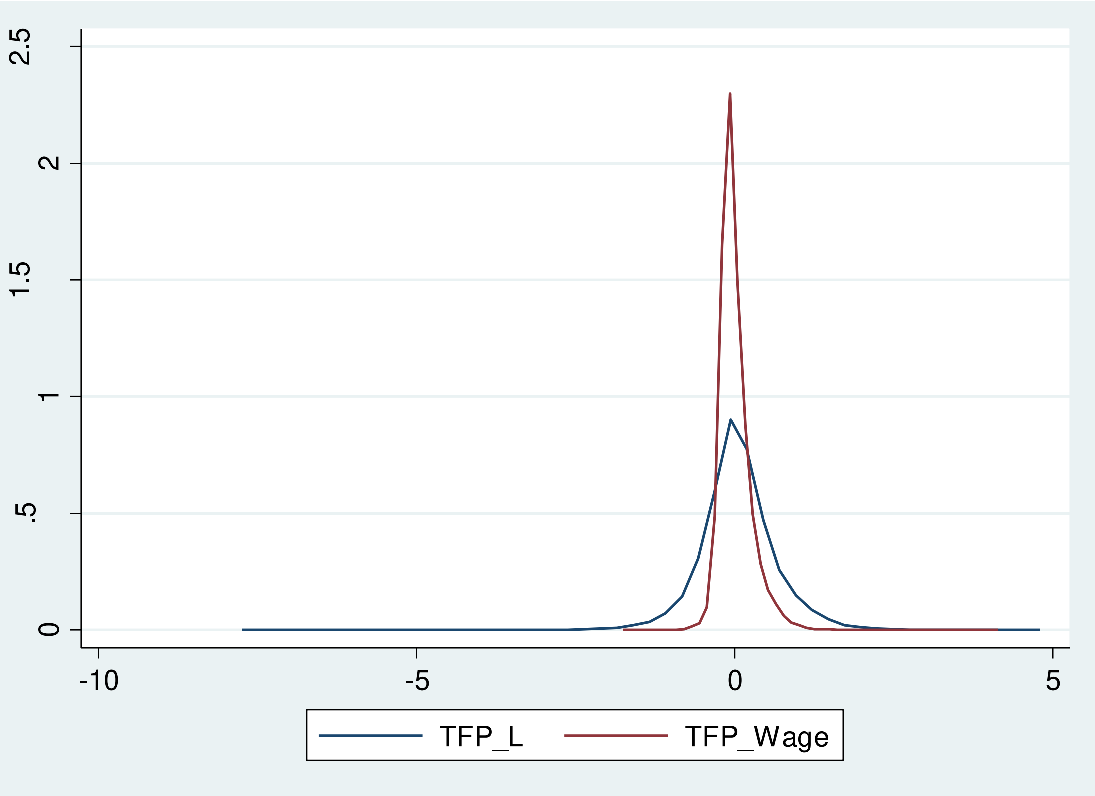
Table 4: T-test for difference in SD of TFP resulted from model with labour count (L) and model with cost of labour input (wage), 2000-2009

Two-digit industry (VSIC1993)	SD mean		T-value
	L	Wage	
Food and drink manufacturing	0.485	0.236	429.4
Tobacco and cigarettes manufacturing	0.609	0.320	49.8
Textile manufacturing	0.483	0.235	336.5
Clothing, leather dying and manufacturing	0.481	0.224	626.1
Leather tanning, fur dressing, luggage and other leather product mfg	0.485	0.249	228.5
Wood processing and wooden product manufacturing	0.481	0.199	531.9
Paper and paper products	0.410	0.149	404.3
Printing, publishing, and recording	0.359	0.126	294.9
Coal, petroleum product and biofuel product production	0.620	0.133	26.1
Chemical and chemical product production	0.652	0.171	400.2
Rubber and plastic product production	0.429	0.169	447.1
Glass, glass product and fine ceramic product production	0.462	0.203	523.4
Metal manufacturing	0.495	0.145	289.4
Metal product manufacturing	0.442	0.166	648.8
Machinery and equipment production	0.465	0.176	180.7
Office and computer equipment production	0.533	0.485	2.34
Electrical and machinery equipment production	0.526	0.182	171.0
Radio, TV and telecoms equipment production	0.676	0.214	112.9
Medical, precious tools & equipment	0.515	0.346	24.1
Vehicle and trailer manufacturing	0.523	0.181	150.9
Other transport equipment manufacturing	0.535	0.212	244.2
Bed, wardrobe, table, chair and other wooden product manufacturing	0.466	0.175	502.2
Recycling	0.402	0.113	41.3
Overall manufacturing	0.475	0.197	1200

Note: TFP and its SD are from FE production function estimates. The production function estimation controlled for time fixed effect.

Table 4 presents results of T-test for the SD differences between the models using two alternative measures of labour input. Overall TFP obtained from models with cost of labour input as a proxy for labour quality input are much less dispersed than that obtained from models with labour input count (see Table 4 and Figure 1). The differences in TFP standard deviations are highly statistically significant except 'Office and computer equipment production' significant at the five per cent level. The evidence suggests the hypothesis that TFP obtained from using cost of labour input is less dispersed than that using labour count is strongly supported.

Figure 1: Kernel distribution of TFP estimated with L and Wage



Note: TFP is estimated by FE estimator for each two-digit industry to allow for different technologies.

The consistent results from the P90/10 ratios in Table 3 and the SD test in Table 4 propose that application of a variable of labour input quality in production function estimation reduces the productivity dispersion remarkably in the case of the Vietnamese manufacturing sector. The wage bill rather than labour conventional variable of labour count better captures the labour input because the VEC does not provide sufficient information on working status: full-time and part-time labour count, under-time and over-time hours worked, while these labour market situations are common in the country. For example, the SOEs may have kept full-time workers but their staffs often underperform, have low productivity and then earn low wage, while private sectors often ask their staffs/workers to work over-time and on weekends to maximize hours of machinery operations in order to minimize cost (depreciation and other fixed unit cost) and to meet their timeline of contract delivery. In such circumstances, labour count is not a proper variable of labour input. Therefore, cost of labour input rather than the conventional measure of labour count should be used in the production function estimation.

5. Some international comparison and explanations to low productivity dispersion in the Vietnamese manufacturing sector

Robust to various production function estimation and techniques used to test the dispersion difference, overall the P90/10 ratio of the manufacturing sector is quite low, 1.78-to-1 for model with wage. [12] This is slightly higher than the UK manufacturing at 1.6-to-1 (Haskel and Martin, 2002), and some transition economies such as Rumania, Hungary, Russia and Ukraine using data in the first half of 2000s (Brown and Earle, 2006), but lower than many other countries including developed countries such as the US, New Zealand, Denmark, and neighbouring developing countries such as Thailand and Malaysia. For example the US manufacturing (Syverson, 2004b) the average P90-10 percentile productivity ratios within industries are over 4-to-1,[13] the New Zealand manufacturing is around 5 and 6-to-1 (Devine et al, 2013), the Danish manufacturing is 3.36-to-1 (Fox and Smeets, 2011), Thailand and Malaysia is about 3-to-1 (The World Bank, 2003).

Given that productivity dispersion is relatively low in Vietnam, but there exists a variation in dispersion across industries. The existing literature suggests that productivity dispersion may be driven by many factors such as firm characteristics, business practices, technology, and competition etc. Amongst others, we consider the role of competition.

We have previously noted that there are some differences in dispersion across sectors and it seems at some extent it relates to competition exposure to each sector. Here we use the most common measure of competition, Herfindahl-Hirschman Index (HHI) to examine how this index is associated with the dispersion level as the literature suggests.[14]

The literature suggests that the dispersion is lower in sectors that have lower HHI indices (greater intensity of competition). Fiercer competition results in lower dispersion because less efficient firms are forced to improve productivity to survive or exit the market if fail to do so (competition in the market). Survival firms'

[12] It is worth noting that as TFP takes into account capital as well a labour, the measures of dispersion derived from TFP will be smaller than those of labour productivity.

[13] The US data in the paper is out-dated back 1977, but other paper (Dhrymes, 1995) used more updated data and finds the ratio is at 2.75. One should bear in mind that the comparison may be invalid due to different time coverage and data cleaning.

[14] We make an assumption that there are no reallocation and selection effects, and hence higher HHI and smaller number of firms reflect lower intensity of competition. The selection and reallocation effects due to competition are well discussed in Boone (2008).

productivity ultimately is likely to be convergent or less dispersed. For example, Oulton (1996) and Haskel and Martin (2002) show that competition plays a role in productivity spread in the UK manufacturing sector. Following the light of these studies, we look at the relationship between competition using HHI and TFP dispersion (P90/10 ratio) through their correlation and a simple regression. The P90/10 ratio and HHI are estimated for each four-digit industry and by year.[15]

First, we observed the correlation between P90/10 and HHI at two digit level is statistically significant at the 1% level (correlation coefficient is positive of 0.62).

Table 5: Regression of dispersion measure (P90/10) on competition measure HHI, dependent variable is P90/10 ratio

Variables	OLS	OLS	OLS	FE	OLS[(a)]	FE[(b)]
HI	1.909**			0.753**	1.909*	0.753*
	[0.136]			[0.115]	[0.967]	[0.355]
Lag of HHI		0.869**				
		[0.132]				
Lag2 of HHI			0.846**			
			[0.166]			
Two-digit dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	1.988**	2.007**	2.049**	2.183**	1.988**	2.183**
	[0.012]	[0.009]	[0.013]	[0.158]	[0.228]	[0.378]
Observations	235,824	148,519	105,842	235,824	235,824	235,824

*Note: Robust (or cluster-bootstrapped for the last two columns) SEs in brackets, +significant at 10%; * significant at 5%; ** significant at 1%. [(a)] and [(b)] is cluster-bootstrapped to correct for spurious significance due to data clustered at four-digit level.*

Second, we regressed the dispersion measure of P90/10 ratio on HHI and control for year and two-digit industry fixed effect. We applied both OLS and fixed effect models, and also did cluster-bootstrapping to correct standard errors due to spurious significance (as we run regression at firm level but data on HHI and dispersion data are at four-digit industry level). Both models OLS and FE of dispersion measure P90/10 on competition measure of HHI is highly significant at the one per cent level (columns 1 to 4, Table 5). We also test the relationship with lagged values (columns 2 and 3). As data is clustered at 4-digit level, hence we run cluster-bootstrapping to correct standard errors. The significance reduced but still significant at the five per cent level (columns 5 and 6, Table 5). The results suggest that markets with higher

[15] P90/10 and SD are of TFP using cost of labour input as labour input in the production function.

HHI or less competition intensity is likely to be more dispersed. In other words, more concentrated industries tend to have higher productivity dispersion. We are not able to make a solid conclusion about causality between competition and dispersion in this paper, but the results at least show a quite clear association between them.

An important question arises here is that why TFP is less dispersed in the Vietnamese manufacturing sector where market competition (or market economy) is not yet to be fully recognized by the Western world? Unlike sparse populated countries where highly dispersed productivity is typically observed, populous countries such as Vietnam may enable firms to learn quicker and allow labour mobility faster. Additionally, competition in Vietnam has increased significantly since the economic reforms in mid 1980s and early 1990s and particularly accession to WTO in 2006 (Doan and Stevens, 2012). The number of firms in the economy increased by more than five-fold from circa 42,000 to about 240,000 firms (Doan and Stevens, 2012) and hence the productivity dispersion ultimately has been narrowed and observed at relatively low level.[16]

The low dispersion in the manufacturing sector in Vietnam may reflect the fiercer competition during the economic transition, but comparing with more mature market economies such as New Zealand, the US, Denmark, Malaysia, Thailand etc. those have higher dispersion suggests that other factors rather than competition may also affect dispersion.

In a case of developed countries, take New Zealand an example, where markets are believed to be highly competitive but the TFP dispersion in manufacturing is still very high (Devine, *et al*, 2013). The possible reason underling this high dispersion would be that the different level-productivity firms co-existed in sparsely geographical locations where competition amongst firms within defined markets may not happen because firms those are mostly small operate in small local markets. Local markets are too small to have many players to compete with. Unlike isolated and small country of New Zealand, firms in Vietnam particularly in the tradable sector of manufacturing are completely exposed not only to domestic competition but also to international competition for domestic markets and for exporting markets from rival

[16] See Appendix 5 for a similar trend of the manufacturing sector

neighbouring countries especially China. This likely helps explain its relatively low productivity dispersion in the Vietnamese manufacturing sector.

6. Summary and conclusions

This paper uses the VEC data and takes into account the effect of labour input quality on productivity dispersion. We estimate production functions for separate industries using both labour count and cost of labour input. We compare the variations in the dispersion across the results to determine whether mismeasurement of labour input increases productivity dispersion within industries. The paper also considers the roles of competition on productivity dispersion. The findings can be summarized as bellow.

First, the labour productivity is highly dispersed and the dispersion varies considerably across industries in the manufacturing sector over the period. Component sectors those are capital intensive are more dispersed than the labour intensive sectors.

Second, TFP dispersion is much lower than labour productivity. It is partly due to the fact that contribution of capital has been taken into account. When including a quality measure of labour input, cost of labour input, reduces remarkably the observed TFP dispersion. Overall we find that including a quality input measure of labour reduces the observed productivity dispersion by 0.15 points or 8.7%. This pattern is observed in almost all industries. Our finding of a considerable decline in dispersion contradicts the existing literature on developed countries (Fox and Smeets, 2011; Bagger *et al*, 2011).

There are remarkable decreases in P90/10 ratios of TFP for non-state sectors or high tech sectors those pay their workers based on individual productivity or performance. Workers' wage hence well captures workers' experience, skills, effort, hours worked and unmeasured individual characteristics such as motivation rather than workers' work presence and their education which is widely believed to be low quality in Vietnam. This may imply that the labour input if mismeasured can result in biases in the calculated dispersion especially for more labour market-oriented industries where the cost of labour input appears to capture price of labour by which workers are paid their marginal value.

Consolidating by further analysis using T-test for the SD differences between the models using two alternative measures of labour input suggests that TFP obtained from models with cost of labour input as a good proxy for labour quality input are much less dispersed than that obtained from models with labour input count. The hypothesis that TFP obtained from using cost of labour input is less dispersed than that using labour count is strongly supported in our study. This also suggests that application of a variable of labour input quality in a production function reduces the productivity dispersion remarkably. The cost of labour input better captures the labour input than conventional variable of labour count because the VEC does not provide detailed information on working status which affects workers' earnings. In such conditions, labour count is not a proper variable of labour input. Therefore, cost of labour input rather than the conventional measure of labour count should be employed in the production function estimation.

Dispersion is relatively low in comparison with some developed as well as developing neighbouring economies. There is some evidence of impact of competition on the dispersion. The increasing competition through the fast growing number of firms over the economic reform forces firms to improve productivity to compete with others to survive and grow. This process would narrow down the productivity gap. Firms in Vietnam are exposed to not only domestic competition but also international competition pressure in domestic markets and in (and for) exporting markets from rival neighbouring countries especially China. This likely helps explain its narrowing productivity dispersion in the Vietnamese manufacturing sector as the literature suggests.

References

Ackerberg, D., Caves, K., & Frazer, G. (2006). *Structural identification of production functions*. Available at http://www.econ.ucla.edu/ackerber/ACF20withtables.html

Bagger, J., Christensen, B. J. and Mortensen, D. (2010). *Wage and productivity dispersion: Labour quality or rent sharing*?, NBER Working Paper.

Bartelsman, E., & Doms, M. (2000). Understanding productivity: Lessons from longitudinal microdata, *Journal of Economic Literature*, *38*(3), 569–594.

Boone, J. (2008). A new way to measure competition. The Economic Journal, 118(August), 1245-1261.

Bloom, N, Schankerman, M., and Van Reenen, J. (2007). *Identifying technology spillovers and product market rivalry.* National Bureau of Economic Research Working Paper 13060.

Brown, D., & Earle, J. (2006). *Employment reallocation, productivity growth, and economic systems: Evidence from transition economies in comparative perspective.* Available at http://www.apeaweb.org/confer/sea06/papers/brown-earle.pdf

Devine, H., Doan, T., & Stevens, P. (2012). *Explaining productivity distribution in New Zealand industries: The effects of input quality on firm productivity differences*. Paper presented at the 2012 NZAE Annual Conference, Palmerston North.

Devine, H., Doan, T., & Stevens, P. (2013). *Explaining productivity distribution in New Zealand industries: The effects of labour input quality on firm productivity differences.* Forthcoming MBIE Working Paper.

Dhrymes, P. (1995). *The Structure of production technology: Productivity and aggregation effects.* London: Edward Elgar.

Faggio, G, Salvanes, K., & Van Reenen, J. (2010). The evolution of inequality in productivity and wages: panel data evidence. *Industrial and Corporate Change, 19*(6): 1919–51.

Foster, L, Haltiwanger, J., & Krizan, C. J. (2006). Market selection, reallocation, and restructuring in the US. retail trade sector in the 1990s. *Review of Economics and Statistics, 88*(4): 748–58.

Foster, L, Haltiwanger, J., & Syverson, C. (2008). Reallocation, firm turnover, and efficiency: selection on productivity or profitability?. *American Economic Review, 98*(1): 394–425.

Fox, J ., & Smeets, V. (2011). *Does input quality drive measured differences in firm productivity?.* National Bureau of Economic Research Working Paper 16853.

Griliches, Z. (1957). Specification bias in estimates of production functions, *Journal of Farm Economics*, *39*(1), 8–20.

Haskel, J., & Martin, R. (2002). *The UK manufacturing productivity spread*, CeRiBA Working Paper.

Hellerstein, J., & Neumark, D. (2006). *Production function and wage equation estimation with heterogeneous labour: evidence from a new matched employer-employee data set*, in “Hard to measure goods and services: Essays in Honour of Zvi Griliches,” University of Chicago, working paper.

Helpman, E. (2006). Trade, FDI, and the organization of firms. *Journal of Economic Literature, 44*(3), 589-630.

Helpman, E., Melitz, M,. & Yeaple, S. (2004). *Export versus FDI with heterogeneous firms,* Harvard University, Department of Economics.

Irarrazabal, A., A. Moxnes, & K.H. Ulltveit-Moe. (2009). *The black box of productivity and the exporter premium*, University of Oslo working paper.

Levisohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservable. *Review of Economic Studies,* 70, 317–341.

Maré, D., & Hyslop, D. (2006). *Worker-firm heterogeneity and matching: An analysis using worker and firm fixed effects estimated from LEED* (November 1, 2006). Statistics New Zealand LEED research paper.

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

Nicoletti, G., & Scarpetta, S. (2005). *Regulation and economic performance: product market reforms and productivity in the OECD.* Organisation for Economic Co-operation and Development Economics Department Working Paper 460.

Olley, G., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry, *Econometrica*, *64* (6), 1263–1297.

Ornaghi, C. (2006). Assessing the effects of measurement errors on the estimation of production functions. *Journal of Applied Econometrics, 21*, 879–891.

Oulton, N. (1996). *Competition and the dispersion of labour productivity amongst UK companies.* National Institute of Economic and Social Research, September 1996, NIESR Discussion Paper No. 103

Syverson, C. (2004a). Market structure and productivity: A concrete example. *Journal of Political Economy, 112*(6): 1181–1222.

Syverson, C. (2004b). Product substitutability and productivity dispersion. *Review of Economics and Statistics, 86*(2): 534–50.

Wooldridge, J. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters, 104*(2009), p.112-114.

World Bank, (2003). *The global economic prospects and the developing countries.* The World Bank, Washington DC, USA.

Appendix 1: Two-digit industries of manufacturing and its description

Code	Industry description
D15	Food and drink manufacturing
D16	Tobacco and cigarettes manufacturing
D17	Textile manufacturing
D18	Clothing, leather dying and manufacturing
D19	Leather tanning and fur dressing, luggage and other leather product mfg
D20	Wood processing and wooden product manufacturing
D21	Paper and paper products
D22	Printing, publishing, and recording
D23	Coal, petroleum product and biofuel product production
D24	Chemical and chemical product production
D25	Rubber and plastic product production
D26	Glass, glass product and fine ceramic product production
D27	Metal manufacturing
D28	Metal product manufacturing
D29	Machinery and equipment production
D30	Office and computer equipment production
D31	Electrical and machinery equipment production
D32	Radio, TV and telecommunications equipment production
D33	Medical & precious tools & equipment
D34	Vehicle and trailer manufacturing
D35	Other transport equipment manufacturing
D36	Bed, wardrobe, table, chair and other wooden product manufacturing
D37	Recycling

Appendix 2: Total Factor Productivity Distribution (Cobb-Douglas OLS) with labour count input vs. cost of labour input

Two-digit industry (VSIC1993)	L		Wage		Difference		%
	P90/10		P90/10		P90/10		
	mean	SD	Mean	SD	Mean	SD	
Food and drink manufacturing	1.97	5.04	1.91	4.82	-0.06	-0.22	-3.1
Tobacco and cigarettes manufacturing	1.41	0.10	1.37	0.09	-0.04	-0.01	-3.0
Textile manufacturing	1.94	0.41	1.82	0.32	-0.12	-0.09	-6.3
Clothing, leather dying and manufacturing	2.04	0.54	1.76	0.70	-0.28	0.16	-13.7
Leather tanning and fur dressing, luggage and other leather product manufacturing	2.26	0.28	1.93	0.37	-0.33	0.09	-14.6
Wood processing and wooden product mfg	1.69	0.10	1.57	0.09	-0.11	-0.01	-6.8
Paper and paper products	1.40	0.08	1.35	0.08	-0.04	-0.01	-3.2
Printing, publishing, and recording	1.43	0.29	1.32	0.24	-0.11	-0.05	-7.6
Coal, petroleum product and biofuel product production	1.61	0.57	1.53	0.53	-0.08	-0.05	-5.2
Chemical and chemical product production	1.53	0.14	1.46	0.14	-0.07	0.00	-4.4
Rubber and plastic product production	1.49	0.11	1.42	0.10	-0.07	-0.01	-4.7
Glass, glass product and fine ceramic product production	1.78	0.19	1.59	0.17	-0.19	-0.02	-10.6
Metal manufacturing	1.36	0.09	1.33	0.11	-0.03	0.01	-2.2
Metal product manufacturing	1.52	0.15	1.43	0.13	-0.09	-0.02	-5.7
Machinery and equipment production	1.54	0.43	1.48	0.58	-0.06	0.15	-4.1
Office and computer equipment production	2.72	3.69	2.81	4.70	0.09	1.01	3.3
Electrical and machinery equipment production	1.54	0.27	1.47	0.25	-0.06	-0.02	-4.2
Radio, TV and telecommunications equipment production	1.99	0.50	1.77	0.35	-0.22	-0.15	-11.3
Medical & precious tools & equipment	2.05	3.42	2.10	5.11	0.05	1.69	2.3
Vehicle and trailer manufacturing	1.67	0.20	1.57	0.22	-0.09	0.03	-5.5
Other transport equipment manufacturing	1.78	0.26	1.65	0.21	-0.13	-0.05	-7.0
Bed, wardrobe, table, chair and other wooden product manufacturing	1.77	3.02	1.60	2.46	-0.17	-0.56	-9.3
Recycling	1.44	0.27	1.34	0.17	-0.10	-0.10	-7.0
Overall manufacturing	**1.74**	**2.51**	**1.63**	**2.39**	**-0.11**	**-0.12**	**-6.6**

Notes: TFP is estimated from models in Table 1. Top one and bottom one percentiles at 4-digit level are removed before computing these numbers.

Appendix 3: Cobb-Douglas Production Function (OLS), pooled 2000-2009

	D15		D16		D17		D18		D19		D20		D21		D22	
Dependent Variable Log VA	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage
lnLN	0.133**	0.139**	0.054*	0.089**	0.233**	0.265**	0.336**	0.376**	0.349**	0.383**	0.176**	0.208**	0.142**	0.167**	0.227**	0.238**
	[0.003]	[0.003]	[0.025]	[0.030]	[0.005]	[0.005]	[0.006]	[0.005]	[0.008]	[0.008]	[0.004]	[0.005]	[0.007]	[0.008]	[0.007]	[0.006]
lnK	0.118**	0.107**	0.089**	0.075**	0.117**	0.092**	0.093**	0.079**	0.089**	0.078**	0.101**	0.088**	0.085**	0.080**	0.082**	0.074**
	[0.002]	[0.002]	[0.018]	[0.019]	[0.004]	[0.004]	[0.005]	[0.004]	[0.006]	[0.005]	[0.006]	[0.003]	[0.005]	[0.004]	[0.004]	[0.003]
lnic	0.749**	0.735**	0.843**	0.809**	0.652**	0.623**	0.579**	0.529**	0.567**	0.525**	0.720**	0.679**	0.776**	0.740**	0.730**	0.687**
	[0.003]	[0.003]	[0.037]	[0.044]	[0.007]	[0.007]	[0.006]	[0.006]	[0.008]	[0.008]	[0.005]	[0.006]	[0.009]	[0.011]	[0.008]	[0.008]
Constant	1.277**	1.152**	0.876**	0.912**	1.532**	1.287**	1.745**	1.234**	1.805**	1.252**	1.294**	1.163**	1.070**	0.945**	1.211**	0.958**
	[0.012]	[0.010]	[0.218]	[0.213]	[0.032]	[0.029]	[0.028]	[0.024]	[0.043]	[0.036]	[0.023]	[0.019]	[0.048]	[0.041]	[0.0375]	[0.032]
Obs	41,411	41,411	233	233	7,992	7,992	13,524	13,524	4,263	4,263	14,137	14,137	7,997	7,997	9,803	9,803
R-squared	0.98	0.98	0.99	0.99	0.96	0.97	0.95	0.96	0.96	0.97	0.97	0.97	0.97	0.98	0.97	0.98

	D23		D24		D25		D26		D27		D28		D29		D30	
	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage
lnL	0.064*	0.064*	0.120**	0.148**	0.166**	0.174**	0.207**	0.269**	0.085**	0.098**	0.167**	0.184**	0.190**	0.208**	0.170**	0.168**
	[0.028]	[0.030]	[0.005]	[0.006]	[0.005]	[0.005]	[0.006]	[0.006]	[0.006]	[0.007]	[0.005]	[0.004]	[0.013]	[0.012]	[0.044]	[0.047]
lnK	0.099**	0.091**	0.108**	0.092**	0.097**	0.086**	0.169**	0.143**	0.080**	0.072**	0.095**	0.084**	0.082**	0.079**	0.039	0.041+
	[0.023]	[0.017]	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.003]	[0.005]	[0.005]	[0.002]	[0.002]	[0.006]	[0.005]	[0.032]	[0.023]
Lnic	0.839**	0.831**	0.797**	0.767**	0.756**	0.736**	0.642**	0.580**	0.840**	0.825**	0.751**	0.725**	0.752**	0.716**	0.843**	0.829**
	[0.024]	[0.030]	[0.006]	[0.006]	[0.007]	[0.007]	[0.006]	[0.007]	[0.008]	[0.009]	[0.004]	[0.005]	[0.011]	[0.013]	[0.035]	[0.049]
Constant	0.796**	0.756**	0.910**	0.792**	1.114**	0.911**	1.400**	1.217**	0.845**	0.761**	1.175**	0.968**	1.117**	0.874**	0.682*	0.377
	[0.132]	[0.124]	[0.027]	[0.024]	[0.032]	[0.028]	[0.020]	[0.016]	[0.046]	[0.043]	[0.025]	[0.022]	[0.050]	[0.042]	[0.265]	[0.237]
Obs	166	166	8,505	8,505	11,647	11,647	15,067	15,067	3,549	3,549	20,360	20,360	5,441	5,441	211	211
R-squared	0.99	0.99	0.98	0.98	0.97	0.97	0.97	0.98	0.99	0.99	0.97	0.97	0.97	0.97	0.96	0.96

	D31		D32		D33		D34		D35		D36		D37	
	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage	Labour count	Wage
lnL	0.120**	0.145**	0.227**	0.283**	0.243**	0.253**	0.200**	0.216**	0.199**	0.219**	0.220**	0.257**	0.107**	0.153**
	[0.011]	[0.011]	[0.015]	[0.019]	[0.047]	[0.073]	[0.014]	[0.013]	[0.012]	[0.010]	[0.005]	[0.005]	[0.023]	[0.018]
lnK	0.093**	0.087**	0.082**	0.059**	0.070**	0.069*	0.100**	0.090**	0.105**	0.099**	0.091**	0.081**	0.111**	0.099**
	[0.008]	[0.008]	[0.012]	[0.010]	[0.024]	[0.031]	[0.009]	[0.008]	[0.008]	[0.006]	[0.003]	[0.002]	[0.020]	[0.018]
Lnic	0.801**	0.771**	0.747**	0.698**	0.732**	0.704**	0.739**	0.706**	0.717**	0.682**	0.697**	0.649**	0.767**	0.740**
	[0.011]	[0.012]	[0.017]	[0.019]	[0.028]	[0.037]	[0.013]	[0.015]	[0.010]	[0.011]	[0.005]	[0.006]	[0.019]	[0.019]
Constant	0.991**	0.837**	1.185**	0.762**	1.223**	0.790**	1.161**	0.926**	1.301**	1.046**	1.309**	1.058**	1.039**	0.955**
	[0.052]	[0.046]	[0.068]	[0.053]	[0.114]	[0.135]	[0.055]	[0.047]	[0.040]	[0.032]	[0.024]	[0.020]	[0.132]	[0.164]
Obs	3,184	3,184	1,751	1,751	787	787	2,545	2,545	4,314	4,314	13,740	13,740	371	371
R-squared	0.98	0.98	0.97	0.97	0.94	0.94	0.98	0.98	0.97	0.97	0.97	0.98	0.98	0.98

*Notes: Robust standard errors in brackets; + significant at 10%; * significant at 5%; ** significant at 1%; All models controlled for year effect. lnL is natural log of labour count or natural log of Wage depending on models with labour count or with cost of labour input. lnK is natural log of cost of capital . To have comparable results, both labour count and wage model use the same sample where both missing log of labour count (L) and missing log of wage are dropped. For industry description, see Appendix 1.*

Appendix 4: T-test for difference in SD of TFP between model with labour count (L) and with cost of labour input (wage) – SD obtained from OLS production function estimation

Two-digit industry (VSIC1993)	2008-2009			2000-2009		
	SD			SD		
	L	Wage	T-val	L	Wage	T-val
Food and drink manufacturing	0.214	0.208	53.7	0.208	0.202	140.6
Tobacco and cigarettes manufacturing	0.070	0.070	3.3	0.122	0.119	9.8
Textile manufacturing	0.211	0.205	24.8	0.230	0.217	60.5
Clothing, leather dying and manufacturing	0.215	0.188	92.8	0.249	0.213	182.7
Leather tanning and fur dressing, luggage and other leather product manufacturing	0.280	0.248	41.6	0.273	0.232	119.2
Wood processing and wooden product mfg	0.194	0.180	64.9	0.188	0.173	124.7
Paper and paper products	0.135	0.126	59.0	0.131	0.122	112.9
Printing, publishing, and recording	0.105	0.089	83.2	0.128	0.112	116.3
Coal, petroleum product and biofuel product production	0.226	0.226	0.0	0.145	0.137	7.0
Chemical and chemical product production	0.142	0.135	25.7	0.161	0.150	77.1
Rubber and plastic product production	0.147	0.133	171.5	0.158	0.147	187.1
Glass, glass product and fine ceramic product production	0.217	0.183	119.9	0.208	0.177	242.2
Metal manufacturing	0.133	0.128	27.2	0.119	0.113	41.6
Metal product manufacturing	0.149	0.138	142.2	0.158	0.145	222.8
Machinery and equipment production	0.165	0.157	22.6	0.163	0.152	45.4
Office and computer equipment production	0.130	0.112	10.0	0.255	0.245	4.1
Electrical and machinery equipment production	0.167	0.160	22.0	0.170	0.158	48.3
Radio, TV and telecommunications equipment production	0.260	0.251	8.9	0.236	0.215	36.0
Medical, precious tools & equipment	0.189	0.190	0.4	0.214	0.202	9.9
Vehicle and trailer manufacturing	0.195	0.185	25.4	0.186	0.171	45.9
Other transport equipment manufacturing	0.215	0.190	29.3	0.208	0.190	53.9
Bed, wardrobe, table, chair and other wooden product manufacturing	0.170	0.150	113.7	0.183	0.161	192.9
Recycling	0.138	0.134	2.6	0.123	0.108	12.4
Overall manufacturing	0.182	0.167	209.3	0.188	0.173	404.5

Appendix 5: Manufacturing sector Firm count and competition over time, 2000-2009

Two-digit industry (VSIC1993)	Firm count			HHI		
	2000/01	2008/09	Diff	2000/01	2008/09	Diff
Food and drink manufacturing	6,934	14,013	7,079	0.117	0.101	-0.017
Tobacco and cigarettes manufacturing	50	51	1	0.127	0.150	0.024
Textile manufacturing	906	3,534	2,628	0.062	0.054	-0.008
Clothing, leather dying and manufacturing	1,370	6,576	5,206	0.032	0.007	-0.025
Leather tanning and fur dressing, luggage and other leather product manufacturing	564	1,740	1,176	0.064	0.051	-0.014
Wood processing and wooden product mfg	1,620	6,453	4,833	0.096	0.024	-0.072
Paper and paper products	884	3,152	2,268	0.136	0.025	-0.111
Printing, publishing, and recording	706	5,405	4,699	0.135	0.048	-0.087
Coal, petroleum product and biofuel product production	24	75	51	0.223	0.525	0.302
Chemical and chemical product production	942	3,546	2,604	0.137	0.199	0.062
Rubber and plastic product production	1,104	5,015	3,911	0.062	0.070	0.007
Glass, glass product and fine ceramic product production	2,285	5,313	3,028	0.203	0.082	-0.121
Metal manufacturing	300	1,534	1,234	0.109	0.066	-0.043
Metal product manufacturing	1,550	10,810	9,260	0.067	0.030	-0.037
Machinery and equipment production	570	2,327	1,757	0.307	0.136	-0.171
Office and computer equipment production	10	151	141	0.994	0.677	-0.317
Electrical and machinery equipment production	364	1,313	949	0.299	0.072	-0.227
Radio, TV and telecommunications equipment production	199	841	642	0.205	0.137	-0.068
Medical, precious tools & equipment	96	344	248	0.482	0.608	0.126
Vehicle and trailer manufacturing	400	742	342	0.347	0.097	-0.250
Other transport equipment manufacturing	585	1,590	1,005	0.199	0.198	-0.001
Bed, wardrobe, table, chair and other wooden product manufacturing	1,263	6,958	5,695	0.140	0.054	-0.086
Recycling	17	213	196	0.768	0.082	-0.685
Overall manufacturing	**22,743**	**81,696**	**58,953**	**0.202**	**0.133**	**-0.069**

Appendix 6: Variable definitions

Key variables used in this paper are defined as follows. Sales of goods and services include total sales of products and services, and other incomes excluding fixed asset sales. Profits are total before-tax profits. Employment comes from counts of employees and working proprietors, an average of year-begin and year-end counts. A working proprietor is assumed to be a person who (i) operates his or her own enterprise or engages independently in a profession or trade, and (ii) receives income from self-employment from which tax is deducted, but not from wages and salary.

Fixed assets are averaged over beginning and ending year fixed assets. Variable costs include intermediate costs (IC) and labour costs. Labour cost includes wages, allowance, contribution to social and health insurance, and union fees paid by firms for employees. The intermediate costs include materials, tools, fuel, electricity, water bills, transport expenses, postage, and insurance. Because IC is not explicitly collected in the census, the IC is estimated as the difference between total sales *minus* sum of labour cost, capital cost (or capital services) and before-tax profits. Capital services cost is estimated as follows:

Capital cost = Depreciation + interest rate*fixed assets

where depreciation is the difference between year-end and year-begin accumulated depreciation. The difference is actually the depreciation incurs during the business year. Some observations with negative depreciation that may be due to selling fixed assets (the difference is negative) were dropped. Interest rate is yearly average interest rate, equals 150% of the base rate of the State Bank of Vietnam (Central Bank). The State Bank of Vietnam periodly sets the base rate for commercial banks, commercial banks are allowed to lend at maximum 150% of the base rate. The commercial banks always lent businesses at 150% of the base rate as the demand for capital in the economy exceeded the capital supply at the 150% of the base rate.