Efficiency or Bounded Rationality?
Drivers of Firm Diversification Strategies in Vietnam

Hien Thu Tran
RMIT International University; 702 Nguyen Van Linh Blvd – Hochiminh city, Vietnam
e-mail: hien.tran@rmit.edu.vn

Enrico Santarelli
University of Bologna, Department of Economics; Piazza Scaravilli, 2 – 40126 Bologna, Italy;
e-mail: enrico.santarelli@unibo.it

Enrico Zaninotto
University of Trento, Department of Economics and Management; Via Inama, 5 – 38122 Trento, Italy;
e-mail: enrico.zaninotto@unitn.it

Abstract

Considering the case of diversified firms within a developing/transition country such as Vietnam, this paper investigates diversification relatedness while taking into account both firm- and industry-specific components. The high volatility of the dynamics of diversification observed in Vietnam suggests the hypothesis that firms decide to enter into new industries following a trial and error process, initiated by boundedly rational herding behaviours, i.e. firms follow the most commonly observed business combinations. Using a survivor-based (SB) measure of relatedness, we test the hypothesis of a boundedly rational behaviour. We find that both the probability of exit and the different performance measures (Return on sales and Total factor productivity) are not or are negatively correlated with SB-related diversification. This is in contrast to what has been observed in developed countries. However, using the SIC distance approach, we obtain the expected positive relationship between performance and relatedness in diversified firms. The conflicting result between these two relatedness indices therefore suggests there has been a trend in follow-up among inexperienced firms that imitate the direction and intensity of the diversification of dominating players within the industry (herd behaviour). However, diversified firms gain experience over time and choose more efficient business combinations in subsequent entries. When we use the classical SIC-based approach, we find that greater diversification raises profitability, but only to an optimum relatedness point, beyond which the positive effect fades away. To control for the endogeneity of diversification relatedness and the serial correlation in error terms, we adopt an instrumental-variable two-stage least-squares estimation approach (IV-2SLS) with GMM treatment.

Keywords: Firm diversification; firm performance; bounded rationality; transition economy.
JEL Codes: L25; L29; P23

Version: 22 March 2015

1. Introduction

Under the assumption that their diversification strategies are driven by efficiency and/or market power motives, firms in developed countries have been shown to exhibit stable behavioural patterns of related diversifications (Palich et al., 2000). The same assumptions and findings do not necessarily apply to developing countries, where firms may act without a planned direction and soon exit recently entered industries. Managerial and technological constraints, as well as low cost of entry, are likely to lead firms to enter and exit new industries, as in a process of opportunity search. One may therefore conjecture that a diversification strategy in developing countries is driven by bounded rationality rather than by efficiency or market power reasons. The identification of what actually motivates firms to diversify into a new industry (either related or unrelated) and of why they soon exit the recently entered industries are therefore tasks important to understanding the overall process of firm and industry dynamics in such countries.

Taking into account both firm-specific and industry-level components, the present study aims to understand the patterns of and rationale behind firm diversification behaviours in Vietnam. We use firm-level data extracted from the annual enterprise survey conducted by the General Statistics Organization (GSO) of Vietnam to disentangle efficiency/market power motives from bounded rationality as drivers of diversification. To this aim, we follow the survivor-based (SB) approach to relatedness in the measurement of how much the frequencies of the actual combinations of four-digit SIC industries deviate from what one would expect if diversification patterns were random (Teece et al., 1994; Piscitello, 2004; Lien and Klein, 2009a and 2013). The results are then compared with those obtained using the SIC distance approaches.

Noteworthy findings include: (i) the SB measure of relatedness is positively associated with the exit decisions of diversified firms, i.e. the more related the diversified industry is to the main industry, the more likely that firms will exit; ii) the significant and positive effect of the interaction term with an industry concentration suggests that the entry decisions of diversified firms follow trends of imitation and boundedly rational herd behaviour; (iii) previous experience with diversification increases the likelihood that firms diversifying in related industries will not exit such industries; (iv) inconsistent with comparable studies (Lien and Klein, 2009a), the SB approach predicts neither the profitability nor the productivity of diversified firms; v) however, when SIC distances are used to analyse the choice of destination industry and its consequent entrepreneurial outcome, we find they are directly

2

correlated to performance (lower probability of exit and higher profitability and productivity), meaning that industrial proximity results in more profitable solutions than the frequency of combinations observed in the market, and firms whose behaviours are guided by technological or product proximities succeed more often than firms that follow the herd.

This paper is structured as follows. Section 2 describes the theoretical framework and develops our hypotheses. Section 3 introduces the empirical strategy and presents the variables adopted. Section 4 provides an overview of the dataset together with descriptive statistics and a pair-wise correlation matrix. Section 5 discusses the estimation results and, finally, section 6 provides some concluding remarks.

2. Theoretical framework and hypotheses

Widely studied with regard to developed countries (Rumelt, 1974; Pennings et al., 1994; Markides, 1995; Coad and Guenther, 2014), diversified firms have received less attention in developing ones (Nachum, 1999 and 2004; Wan, 2005), where neglected is in particular the issue of diversification strategies pursued by small firms. There has been a common pathway among Asian developing countries, particularly China, India, South Korea and Vietnam, which transformed state-owned firms into large diversified firms as the main route of growth (Economist, 1997a, 1997b, 1997c; Loc et al. 2006). However, recent empirical findings show that in Vietnam, smaller (and younger) diversified entrepreneurial firms are more successful than their larger counterparts (Santarelli and Tran, 2013).

Diversification decisions can be driven either by efficiency/market power motives or by opportunism (Nachum, 1999; Montgomery, 1985; Dawid and Reimann, 2011). Transaction costs economics claims that the efficiency calculus of the neoclassical theory ignores possible opportunistic behaviours, whereas the resource-based theory suggests that it underestimates input indivisibility and factor market imperfections (Wernerfelt, 1984; Ng, 2007). Whenever environmental uncertainty raises the likelihood of undesirable outcomes, imitative behaviours can drive diversification, as following their predecessors is less risky for firms. This *herd behaviour* is scrutinised in the relevant literature from both economic and social perspectives.

From an economic perspective, an imitative behaviour can be justified by either information or rivalry reasons (Lieberman and Asaba, 2006). Within the domain of the information approach, the leading theory of herd behaviour is called *information cascades* (Banerjee, 1992; Bernardo and Welch, 2001). Information cascades occur when managers

are unable to assess the connection between actions and outcomes (Morone, 2012). Under such circumstances, it is thus "optimal for an individual, having observed the actions of those ahead of him, to follow the behaviour of the preceding individual without regard to his own information" (Bikhchandani et al., 1992: 992). Empirically, "rational" herding occurs when decision makers suppress their private information, either because making a bad decision is less costly when others make the same decision (Scharfstein and Stein, 1990) or because they believe the decisions of others reflect valuable private information (Banerjee, 1992; Bikhchandani et al., 1992). In the same fashion, when network externalities give rise to industry standards, firms imitate to minimise costs (Katz and Shapiro, 1985) or follow successful "first movers" to extract the beneficial spillover (Lieberman and Montgomery, 1988).

A different rationale for imitative diversification comes from rivalry theories. Firms with comparable resource endowments and market shares can or cannot pursue diversification strategies (Lieberman and Asaba, 2006). As diversification strategies are often difficult and risky (Gimeno and Chen, 1998) to pursue, firms increasingly adopt homogenous strategies to maintain their relative position and to neutralise the aggressive actions of rivals. Particularly, frequent contacts across markets allow firms to respond to aggressive actions from their multi-market rivals in other markets, and the threat of such retaliation eases the intensity of market competition in the focal market (Karnani and Wernerfelt, 1985). This type of imitative behaviour is underlined by the mutual forbearance hypothesis, first proposed by Edwards (1955) and empirically supported by subsequent studies (e.g. Greve and Baum, 2001), which posits that firms prefer entering industries in which they will meet existing competitors as a means of establishing mutual forbearance. However, mutual forbearance should be associated with a minimum level of concentration in relevant markets to be a plausible motive for portfolio choices.

Economic reasons for herding behaviours can be fostered by social phenomena: some managers may deliberately imitate the decisions of their peers to avoid a negative reputation and may therefore enhance their status (Palley, 1995; Scharfstein and Stein, 1990). According to the legitimacy theory (DiMaggio and Powell, 1991), institutional pressures for social conformity enhance homogeneity or isomorphism among firms exposed to public scrutiny or government control, and firms seek legitimacy in the eyes of important constituents and stakeholders by adopting structures that are considered appropriate and

rational. Moreover, it has been shown that larger organisations are more likely to be imitated (Haunschild and Miner, 1997).

In uncertain environments, as is the case with developing countries, imitative behaviours can produce mistakes that result in early exits and unprofitable business combinations. The observation of the successful actions of others may raise one's aspiration levels beyond what can realistically be attained (Greve, 1998; Narduzzo and Warglien, 1996). Besides, an imitative behaviour can lead to underestimating the effort and resources needed to achieve a successful result (Westphal et al., 1997; Fligstein, 1985). If the wrong path is chosen, imitation can be costly for firms and for society (Lieberman and Asaba, 2006). On the other hand, the low costs of entry and exit suggest a trial and error procedure, activated by imitation, through which firms iteratively search for the best combination of businesses.

In conclusion, the herding behaviours of diversified firms might result from either rational imitation for both efficiency/market motives (mutual forbearance) or bounded rationality, induced by managerial and technological constraints and the low costs of entry and exit. Thus, this paper will investigate diversification relatedness by contrasting two alternative hypotheses:

H1: Firm diversification is driven by rational herding: the greater the relatedness of the newly entered industry to the currently observed business portfolio, the less likely it is the firm will exit the newly entered business, other things equal.

H2: Firm diversification is driven by boundedly rational herding: the greater the relatedness of the newly entered industry to the currently observed business portfolio, the more likely it is the firm will exit the newly entered business, other things equal.

In order to assess our hypotheses, we consider an index of market relatedness as an indicator of imitative behaviour. Relatedness primarily occurs at the inter-industry level. Research on corporate diversification in developed countries has repeatedly documented the existence of stable and systematic patterns in diversification strategies that are not firm-specific (Silverman, 1999). For instance, some industry combinations are apparently perceived by decision makers as more attractive than others are (Chatterjee and Wernerfelt, 1991; Lien and Klein, 2009a; Teece et al., 1994). In order to check for the presence of systematic business combinations, an SB index of diversification, which measures the observed frequencies of business combinations with respect to a random pattern of diversification, is used. This index has been adopted in the past (Teece et al., 1994; Lien and Klein, 2009a) as an indicator of the coherence of corporate businesses with respect to the empirical combination of businesses more frequently observed in the market; and a

significant relationship between relatedness (businesses coherence) and performance was proved. In particular, Lien and Klein (2009b) suggest that relatedness, measured as observed pattern of industry combinations, drives entry and makes exit less probable. In our setting, instead, the SB relatedness of a corporate business portfolio can simply result from imitative behaviour, reflecting a temporary combination reached during the trial and error process due to boundedly rational herd behaviour. If this was the case, we should not observe a correlation between SB relatedness and firm performance. To complete our test, we consider a different measure of relatedness, the more common SIC index, which measures the distance between industries in the classification tree. Despite some arbitrariness of industrial classification, SIC code proximity reflects a kind of technological or product proximity and can be considered a reliable measure of the economic relatedness between businesses. The comparison between the SB and SIC indices of relatedness is then used to gather hints as to the role of imitative versus economic motives of business pairing inside a firm. Finally, we control for the hypothesis that rational herd behaviours are driven by mutual forbearance.

3. The empirical model

We use two models to test our hypotheses. The first is an exit model: the high volatility of diversification decisions suggests we determine whether an exit from an industry can be positively related to the business pairing in the most frequent combinations observed in the market. The second model uses two different measures of performance: return on sales (ROS) and total factor productivity (TFP). In both models, the effect of related diversification is measured controlling for the sensitivity of performance both on industrial structure and firm-specific phenomena, in line with a widespread approach in empirical industrial economics (Martin, 2002 and, as regards diversification, Christensen and Montgomery, 1981) that joins the traditional Structure-Conduct-Performance (SCP) paradigm with the Resource Based View (RBV).

3.1 Exit

Rather than testing the potentiality of diversifying for the entire set of industries that were not entered by our diversifying firms, we instead look at the probability of exiting the recently entered industry. It is our conjecture that the diversification behaviours of firms in Vietnam follow the pressures of conformity with dominant players in their business network (herd behaviour). This is especially prevalent in situations where entrepreneurs have little

business experience and believe it is less risky to follow the common diversification direction of the group. As a result, entrepreneurs following the herd will be more likely to exit the newly diversified industries than their counterparts who based their diversification decisions on rational factors. The general model is thus:

$$P(exit = 1) = \alpha_1 + \alpha_2 Ind_i + \alpha_3 Ind_j + \alpha_4 Fir_{kt} + \alpha_5 SB_i (or\ \alpha_5 SIC_i) + \varepsilon_{kt} \ (1)$$

Where $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$: estimated parameters; Ind_i, Ind_j: industry-level characteristics of the firm's target industry i and main industry j, respectively; Fir_{kt}: firm-level characteristics of firm k at time t; SB_i: sales-weighted average SB relatedness of the target industry i to all other industries in the portfolio of the firm k; SIC_i: sales-weighted average SIC-based relatedness of the target industry i to all other industries in the portfolio of the firm k; ε_{kt}: error terms.

It is crucial to note that because exits—along with non-entry—determine the key independent relatedness variable, and our dependent variable is the probability of exit, endogeneity may be a potential concern. However, what we are investigating is essentially how much information non-entry and exit decisions by other firms at time t_0 contain about the probability of a given firm exiting a given industry by time t_1. Thus, past exit decisions cannot affect future exit decisions because firms cannot exit the same industry twice; alternatively, past decisions against entering an industry cannot affect future exit decisions because a firm cannot exit an industry that it has not entered. We adopt the random logit model to estimate this equation.

3.2 Performance

In order to detect whether diversified firms have superior performance (profitability, ROS or TFP) compared to their non-diversified counterparts, we estimate the following model:

$$ROS_{kt}(TFP_{kt}) = \beta_1 + SB_{jt}\beta_2 (or\ SIC_{jt}\beta_2) + Fir_{kt}\beta_3 + Ind_j\beta_4 + \epsilon_{kt} \ (2)$$

Where $\beta_1, \beta_2, \beta_3, \beta_4$: estimated parameters; Ind_j: industry-level characteristics of the firm's main industry j; Fir_{kt}: firm-level characteristics of firm k at time t; SB_j: sales-weighted average SB relatedness of the main industry j to all other industries in the portfolio of the firm k; SIC_j: sales-weighted average SIC-based relatedness of the main industry j to all other industries in the portfolio of the firm k; ϵ_{kt}: error terms.

At this point, we are ready to check for possible violations of the underlying estimation assumptions in order to choose the most appropriate estimation method.

Test for violations of estimation assumptions: $E(\epsilon_{kt}\epsilon_{k't'}) = \begin{cases} \sigma_\epsilon^2 & k = k', t = t' \quad (H1) \\ 0 & otherwise \quad (H2) \end{cases}$

- Heteroskedasticity (H1): We apply the White test for heteroskedasticity to the panel data to determine the strong existence of heteroskedasticity in our data[1].

- Serial correlation in the time-series data (H2): the Wooldridge test for the first-order autocorrelation in the panel data is significant even at the 1% level, which indicates the presence of a first-order serial correlation for both the ROS and TFP equation[2].

- Endogeneity of diversification index: The Durbin-Wu-Hausman test does indicate a strong presence of the endogeneity of diversification for both relatedness measures[3].

Thus, several econometric problems arise from our estimating equation (2): (i) Relatedness index SB_{jt} and SIC_{jt} are found to be endogenous; (ii) First-order serial correlation is present in the error terms; (iii) The panel dataset has a short time dimension ($T = 6$) and a large number of firms ($n = 14212$). To deal with these problems, we apply two approaches: (i) the Prais-Winsten and Cochrane-Orcutt (prais) generalised least-squares methods in which errors follow a first-order serial correlation and (ii) the IV-2SLS with GMM treatment to control for the endogeneity of relatedness measures (ivreg2) and clustering across firms, which is efficient in the presence of endogeneity and a first-order serial correlation (Baum and Schaffer, 2003). The IV-GMM treatment requires the availability and validity of exogenous instruments that are correlated with diversification intensity, but that are uncorrelated with firm performance (ROS/TFP). In this paper, we use industry-size dispersion (proxy for industry life-cycle stage) and the industry concentration ratio as the IVs. According to Hu et al. (2005), industry-level variables could potentially become effective instruments to correct for firm-specific effects, as these variables define the environment in which the firms operate and yet are independent of a firm's specific characteristics.

---

[1] White test for the presence of heteroskedasticity: $\chi^2(115) = 1649$; p-value=0.000
[2] Test for serial correlation: $F_{ROS}(1, 11557) = 7.952$, p-value = 0.0048; $F_{TFP}(1, 11557) = 10.568$, p-value = 0.000
[3] Durbin-Wu-Hausman test: SIC-based index $\chi^2(1) = 6.755$; p-value = 0.00935; SB index: $\chi^2(1) = 10.267$; p-value = 0.00135

Appendix 1 describes all the variables that we use for the analysis: firm performance variables, industry-level and firm-level control variables, and diversification relatedness measures. Appendix 2 reviews different diversification measures which have so far been used in the literature; while appendix 3 gives a detailed construction of our survivor-based index of relatedness.

4   Data description

Panel firm-level data from 2000 to 2005 were extracted from the annual national enterprise surveys conducted by Vietnam's GSO, providing data shown to be particularly appropriated for the study of firm dynamics (cf., among others, Santarelli and Tran, 2012; Ha and Kyota, 2014). The dataset allows us to observe the diversification patterns of firms over time and track their behaviours, as well as to determine how these behaviours influence their performance. The survey is comprehensive and harmonised across provinces and industries to obtain a coherent view of the various aspects of firms, including segment data (ISIC code, industry sales, size and assets), accounting data (debts, revenue, profit, assets), basic demographic data (year of inception, ownership type, size of labour force) and innovation data (R&D expenditure, innovation investment, technical personnel). We use the whole population of diversified firms in all 64 provinces of Vietnam. After dropping outliers and missing values in the variables of interest, we end up with an unbalanced panel of 67,522 observations, corresponding to 5,333 firms in 2000; 9,991 firms in 2001; 11,174 firms in 2002; 12,742 firms in 2003; 14,070 firms in 2004 and 14,212 firms in 2005. These firms are active in 502 different industries. Of the 125,751 possible industry pairs, 9,274 pairs are observed.

Table 1 presents the number of firms initiating and stopping diversification each year[4]. Diversification decisions are unstable and decay quickly over time; many firms attempt to diversify, but they abandon their strategy immediately because they receive unexpected results. In fact, there are only 2,034 consistently diversified firms (around 11%), which started diversifying before 2000 and still persistently continued to diversify during the subsequent years. Many firms started their diversification activity before or during 2000, stopped for a while and then restarted. A plausible explanation of this pattern of behaviour is that these firms decided to cease their unprofitable diversifications and they determined a

---

[4] The detailed list of diversified firms could be obtained from the authors.

new potentially successful diversification direction. Therefore, they restarted their diversification strategy the following year.

Table 1: Diversification dynamics in each year[5]

Year	Total firms	Diversified firms	Undiversified firms	Initiate diversification	Stop diversification
2000	40407 (100%)	3077 (7.6%)	37330 (92.4%)		
2001	53247 (100%)	4720 (8.9%)	48527 (91.1%)	3018 (5.7%)	1375 (2.6%)
2002	60932 (100%)	3973 (6.5%)	56959 ((93.5%)	1480 (2.4%)	2227 (3.6%)
2003	69631 (100%)	6387 (9.2%)	63244 (90.8%)	3758 (5.4%)	1344 (1.9%)
2004	88474 (100%)	7846 (8.9%)	80628 (91.1%)	3919 (4.4%)	2460 (2.8%)
2005	105811 (100%)	8622 (8.1%)	97189 (91.9%)	4075 (3.8%)	3299 (3.1%)

Figure 1 shows the percentage of diversified firms by agriculture/ manufacturing/ service industries, respectively, over the years. In general, while diversified firms in the agriculture sector remain low consistently (less than 1%), diversification activities are quite fairly distributed between manufacturing and service firms. Figure 2 presents the pattern of the diversification behaviours of firms that diversify for the first time. In general, around half of them diversify within the same industry. It is obviously less risky to enter a similar business to capitalise one's mainstream competencies. The remaining half of the firms attempt to provide complementary business activities by diversifying into a completely new business sector. Around 30% of firms enter the service sector; nearly 20% enter the manufacturing sector and only a few enter the agriculture sector. Table 2 presents the tabulation of firm age and ownership type: 78% of diversified firms are less than 10 years old, and a private firm is the most common ownership type among diversified firms. Summary statistics and the correlation coefficients of all variables are presented in Appendix 4. We assume that the relatedness between industries as measured by SR_{ij} changes a little over the relevant period. The primary measure of SR_{ij} is calculated for the year 2000, using all firms active in two or more four-digit SIC codes as a basis. For those industry pairs that were not combined by 2000, we calculate their SR_{ij} using the data of subsequent periods.

---

[5] As the numbers of diversified firms reported in Table 1 do not include those firms exiting diversification temporarily but initiating diversification soon thereafter, they are smaller than those reported above.

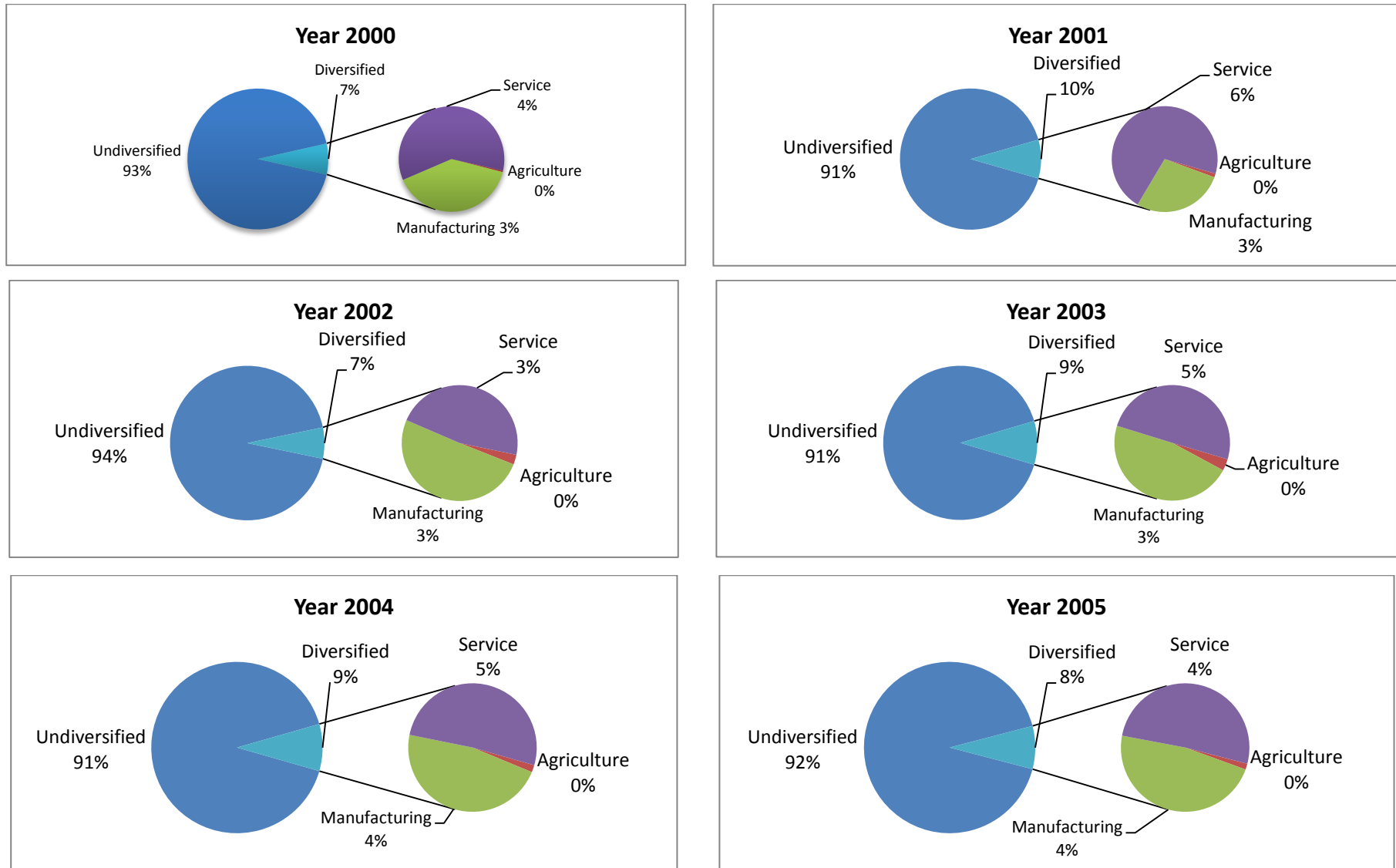**Figure 1: Tabulation of diversified firms by industries from 2000 to 2005**

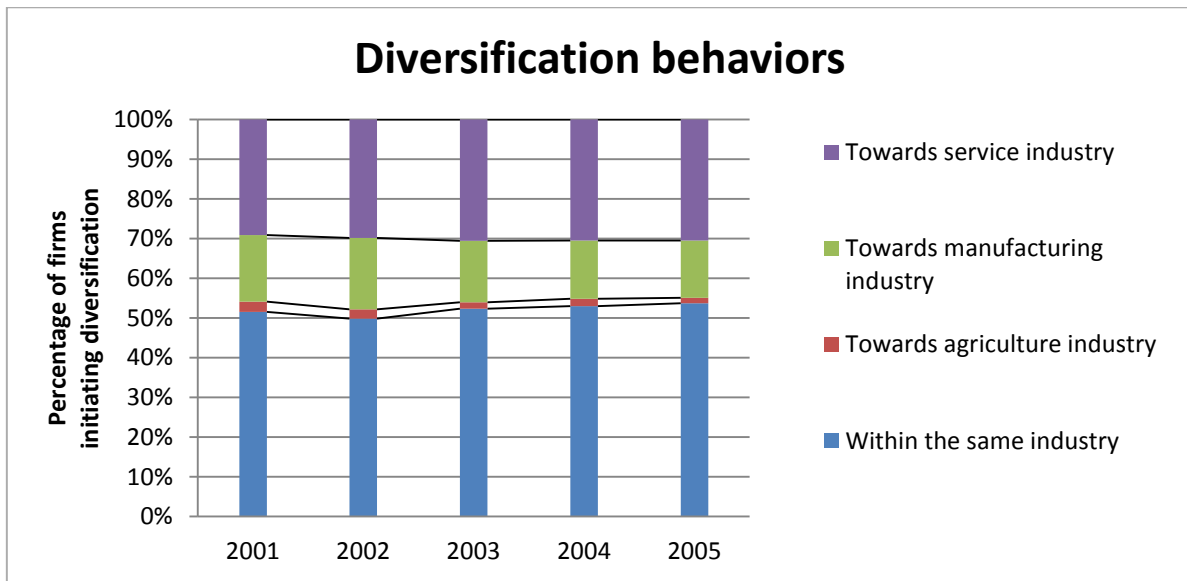**Figure 2: Dynamics of diversification behaviours**

Table 2: Tabulation of firm age and ownership types of diversified firms in 2005

Firm age	State-owned	Cooperative/Partnership	Limited liability	Private	Cooperation	Foreign-invested	Total
0 – 10 yrs	722	569	3675	5264	1474	111	11815 (78)
11 – 20 yrs	729	109	515	741	259	91	2444 (16.1)
21 – 30 yrs	193	64	36	96	79	16	484 (3.2)
31 – 40 yrs	100	11	10	40	47	5	213 (1.4)
41 – 50 yrs	113	35	9	19	33	2	211 (1.4)
> 50 yrs	8	1	3	1	3	1	17 (0.1)
Total	1865(12.3)	789 (5.2)	4248 (28)	6161 (40.6)	1895 (12.5)	226 (1.5)	15184 (100)

5   Empirical results

5.1 Probability of exiting the recently entered industry

The results from random logit regressions are shown in Table 3. The first three regressions (column 1, 2, 3) present the estimation results of the probability of exiting the first-, second- and third-entered industries using the SB relatedness measure. The next three regressions (column 4, 5, 6) present the comparable results using an SIC-based relatedness measure.

If the SB measure was related to rational herding, it should be negatively related to the probability of exit. However, we do not observe that pattern here when the SB relatedness measure is used. As seen from regression 1 in Table 3, the coefficient is positive but not statistically significant. Thus, either the competitive process does not filter inefficient firms

12

out of the market or the traditional efficiency reason is not the main motivation for firm-level diversification. However, from the results in the next two regressions (column 2 and 3), we can see the effect of experience: the higher the relatedness of the second- and third-entered industries with the diversified firm's current portfolio of business, the less likely it is the firm will exit that industry. The effect of experience can be seen from the increasing statistical and numerical significance of SB-relatedness coefficients from 0.002 to 0.019. It is our conjecture that in their first attempt, single business firms, due to bounded rationality, may decide to diversify not for traditional efficiency reasons; for instance, following successful network partners (herd behaviour), group peer pressure, etc. The failure of the first entry and being swept out of the market due to the competitive process will provide them with experience to make smarter choices of more related industries for the next diversification. This may suggest an information cascade model through which firms learn to use signals resulting from others' decisions, and they balance them with private information (pertaining, for instance, to the resources needed to manage diversified activities). Therefore, we reject the former explanation that the competitive process does not do a good job in filtering inefficient firms out of the market; and we doubt that the first diversification of a firm is induced by boundedly rational herding, which partially supports Hypothesis 2 and contradicts Hypothesis 1. On the other hand, when an SIC relatedness measure is used, the role of the market or technical proximity in the context of diversification strongly emerges. The SIC approach is based on standardised distances in the SIC system, and it is little influenced by the behaviours of other firms. The coefficients increase from -0.293 to -0.93 and are statistically significant at the 1 percent level. In conclusion, firms entering SIC-related industries are more likely to survive and foster learning from their experience.

In the next two regressions (column 7 and 8) of Table 3, we address the possibility that herd behaviours reflect mutual forbearance rather than efficiency. According to Lien and Klein (2009a), the interaction between concentration and relatedness can suggest a mutual forbearance behaviour that may decrease the probability of exiting. Not only we do not observe this effect, but the cross term is negatively related to business survival. This result can be interpreted in light of boundedly rational herd behaviours: far from imitating major firms for mutual forbearance, firms simply mimic the combination of businesses in concentrated industries, because they give more weight to the observation of relatively larger firms. Once again, the effect of the SIC relatedness measure contradicts that of the SB measure. Firms choosing SB-efficient combinations will be more likely to fail and exit,

whereas those entering SIC-related industries will survive. We again have statistical evidence to reject Hypothesis 1 in favour of Hypothesis 2: if the firm's diversification is driven by boundedly rational herding, the greater the relatedness of the newly entered industry with its current portfolio of businesses, the more likely it will be that the firm exits that industry, other things equal.

As the social capital from weak ties, such as business networks, business associations and groups, is strong and pervasive, firms perceive that it may be less risky to follow their network partners or their competitors in entering a new market (Santarelli and Tran, 2013). This herding occurs when decision makers suppress their private information, either because making a bad decision is less costly when others make the same decision (Scharfstein and Stein, 1990) or because decision makers believe the decisions of others reflect valuable private information (Banerjee, 1992; Bikhchandani et al., 1992). Either way, entry decisions may be based on the actions of others rather than on superior private knowledge about which industries are related to each other. However, once entry has occurred, competitive forces begin to screen the good decisions from the bad, which should result in the poor performances of bad decision makers.

In the final two regressions (column 9 and 10), we control for the effect of total factor productivity (TFP) on the likelihood of exiting the new industry, and we explore the indirect effect of productivity through relatedness by adding the interaction term between SB/SIC relatedness indices and TFP.

The latter attempt is to determine whether firms that are more productive are likelier to survive in the newly entered industry than less productive ones, assuming they have the same relatedness profile. When controlling for total factor productivity and its interaction with the relatedness indices, all estimated parameters are consistent with earlier findings, reaffirming our interpretation: the first diversification of a firm is induced by boundedly rational herding rather than by efficiency motives. As expected, the negative and significant coefficients of productivity indicate that firms that are more productive are less likely to exit the new industry. This positive effect of productivity is much stronger when we consider the interaction between relatedness and productivity. Given the same relatedness profile, firms with higher productivity will be more likely to survive than their counterparts. Productivity increases significantly a firm's survival propensity in their diversification attempts.

Table 3: Probability of exiting a recently entered industry

Variables	Random Logit Regression Output on the Probability of Exit									
	Survivor-based relatedness measure			SIC-relatedness measure			Interaction between ind. concentration and relatedness		Interaction between productivity and relatedness	
	Industry 1[a] (1)	Industry 2[b] (2)	Industry 3[c] (3)	Industry 1[a] (4)	Industry 2[b] (5)	Industry 3[c] (6)	SB (7)	SIC (8)	SB (9)	SIC (10)
Survivor-based relatedness	0.002	-.018**	-0.019*				0.026*		0.02**	
	(0.002)	(0.003)	(0.011)				(0.003)		(0.002)	
SIC-based relatedness				-0.293**	-0.797**	-0.93**		-0.369**		-0.349**
				(0.047)	(0.175)	(0.36)		(0.085)		(0.05)
Main industry Size dispersion	0.000	-0.000	-0.000	0.000**	-0.000	0.000	0.000	0.000**	0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Main industry ROA	-0.239*	0.559	1.15	-0.135	0.448	1.155	-0.234*	-0.122	-0.227*	-0.121
	(0.130)	(0.467)	(0.794)	(0.131)	(0.427)	(0.802)	(0.131)	(0.133)	(0.13)	(0.132)
Entered industry Size dispersion	0.000**	0.000**	0.000	0.000**	0.000*	0.000	0.000**	0.000**	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Entered industry ROA	-1.243**	-1.309**	-1.835*	-0.994**	-1.321**	-1.572*	-1.331**	-1.132**	-1.197**	-0.929**
	(0.241)	(0.513)	(0.869)	(0.235)	(0.515)	(0.852)	(0.244)	(0.241)	(0.240)	(0.234)
Capital intensity	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Debt ratio	-0.001	-0.023	-0.032	-0.001	-0.021	-0.036	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.027)	(0.065)	(0.001)	(0.027)	(0.065)	(0.001)	(0.001)	(0.001)	(0.001)
Innovation intensity	2.24**	2.127**	3.11*	2.243**	2.065**	3.11*	2.217**	2.215**	2.232**	2.238**
	(0.217)	(0.614)	(1.35)	(0.217)	(0.608)	(1.34)	(0.217)	(0.217)	(0.217)	(0.217)
Innovation intensity squared	-3.598**	-4.12**	-5.81*	-3.592**	-4.031**	-5.799*	-3.583**	-3.573**	-3.602**	-3.601**
	(0.286)	(0.976)	(2.68)	(0.286)	(0.965)	(2.66)	(0.286)	(0.286)	(0.286)	(0.286)
Export	0.201**	0.011	0.022	0.209**	0.028	0.011	0.197**	0.202	0.217**	0.225**
	(0.041)	(0.095)	(0.157)	(0.041)	(0.094)	(0.156)	(0.041)	(0.041)	(0.041)	(0.041)
Firm age	0.021**	0.001	-0.004	0.021**	0.002	-0.004	0.021**	0.021**	0.021**	0.021**
	(0.002)	(0.004)	(0.007)	(0.002)	(0.004)	(0.007)	(0.002)	(0.002)	(0.002)	(0.002)
Labour size	0.638**	0.814**	0.496	0.617**	0.779**	0.401	0.637**	0.618**	0.610**	0.59**
	(0.061)	(0.191)	(0.445)	(0.061)	(0.191)	(0.441)	(0.061)	(0.061)	(0.061)	(0.061)
Labour size squared	-0.071**	-0.063**	-0.025	-0.071**	-0.063**	-0.017	-0.071**	-0.071**	-0.072**	-0.072**
	(0.007)	(0.018)	(0.037)	(0.007)	(0.018)	(0.037)	(0.007)	(0.007)	(0.007)	(0.007)
Economic size	-0.029	-0.163	1.107*	-0.008	0.164	1.135*	-0.032	-0.022	-0.041	-0.019

	(0.077)	(0.229)	(0.556)	(0.078)	(0.229)	(0.555)	(0.078)	(0.078)	(0.078)	(0.078)
Economic size squared	0.011**	0.002	-0.049*	0.011*	-0.002	-0.051*	0.011**	0.011**	0.014**	0.013**
	(0.004)	(0.011)	(0.025)	(0.004)	(0.011)	(0.025)	(0.004)	(0.004)	(0.004)	(0.004)
Industry Concentration							0.463**	0.559**		
							(0.099)	(0.094)		
Concentration* relatedness							0.025**	-0.266		
							(0.008)	(0.23)		
TFP									-0.000*	-0.000*
									(0.0000)	(0.0000)
TFP * relatedness									-0.000**	-0.000*
									(0.000)	(0.0000)
Ownership type dummies $\chi^2(5)$	191.6**	5.86	7.53	231.8**	5.75	6.61	172.5**	199.8**	197.3**	255.3**
Intercept	-4.14**	-5.65**	-9.33**	-4.298**	-5.539**	-9.35**	-3.984**	-4.031**	-4.841**	-5.312**
	(0.313)	(0.978)	(2.46)	(0.314)	(0.977)	(2.46)	(0.316)	(0.318)	(0.343)	(0.343)
Wald statistic	$\chi^2(20)$ 1680.8**	$\chi^2(20)$ 180.13**	$\chi^2(20)$ 58.66**	$\chi^2(20)$ 1631.2**	$\chi^2(20)$ 177.46**	$\chi^2(20)$ 62**	$\chi^2(22)$ 1695.5**	$\chi^2(22)$ 1650.1**	$\chi^2(22)$ 1695.7**	$\chi^2(22)$ 1653.36**
Observations	30823	6381	1752	30823	6381	1752	30820	30820	30820	30820

Note: [a]: the probability of exiting the first-entered industry; [b]: the probability of exiting the second-entered industry; [a]: the probability of exiting the third-entered industry
*: significant at the 5% level; **: significant at the 1% level. Standard errors in parentheses
We control for provincial fixed effects by three dummies *south, central* and *north* for isolating those provinces located in the North, Centre and South of Vietnam, respectively[6].

---

[6] We initially included 64 dummies to control for the fixed-effects of 64 provinces in Vietnam, but this caused an overfitting model and collinearity problems.

Other noteworthy findings include: (i) industry profitability is negatively associated with the exit rate of diversified firms. Obviously, firms will be less likely to fail when their main industry and/or the recently entered industry are still able to offer an abundance of profitable opportunities; (ii) highly concentrated industries filter entrants in a significantly faster manner; (iii) it is tougher for technologically intensive firms to survive in new industries. Innovation investments entail high sunk costs and asset specificity, making transferring and deploying to other new industries difficult; and (iv) diversification seems to be a prominent activity among small and young firms. Firms diversifying after birth are more likely to survive than their older and larger counterparts at the moment of diversification.

5.2 Performance of diversified firms

We assume that the ability to survive in the newly entered industries is closely related to performance, as well as either profitability (ROS) or productivity (TFP). Table 4 presents in the first four regressions (1- 4) the results of the *profitability* equation controlling for the relatedness of firms' portfolios of businesses with respect to their main industry and other relevant industry- and firm-specific characteristics. Due to the endogeneity of diversification relatedness, the IV-GMM model provides more unbiased and efficient estimations, based on which our results will be discussed. The Hansen test of over-identifying restrictions supports the null hypothesis that the instruments are valid.

With ROS as a measure of profitability, we find contradictory results between the SB relatedness and the SIC relatedness equations. While the coefficient of SB relatedness is weakly and insignificantly related to firm profitability, we find a statistically significant positive relationship between SIC relatedness and firm performance. This result is consistent with the above findings of the herding incentives among diversified firms. Firms lacking knowledge and experience in judging the efficiency of their diversification decisions will follow other firms in choosing the industry to enter. This information cascade (or crowd effect) on one hand is assumed to be less risky and safer for young and small firms; on the other hand, it results in "bad entries" and the presence of under-performing firms (Santarelli and Vivarelli, 2007). The emerging market in Vietnam is dynamic and competitive, such that the "revolving door" mechanism is so efficient that a number of new entries in the previous year will immediately cause an equivalent number of exits from both unprofitable incumbents

and entry mistakes (Santarelli and Tran, 2012). Thus, we have enough evidences to claim that those firms following the herd will perform less efficiently than their counterparts.

Table 4: Profitability of diversified firms

	Profitability (ROS)				Productivity (TFP)			
	SB relatedness		SIC relatedness		SB relatedness		SIC relatedness	
	GLS[1]	IV-GMM[2]	GLS[1]	IV-GMM[2]	GLS[1]	IV-GMM[2]	GLS[1]	IV-GMM[2]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SB-index	0.0002	0.0001			0.0005	0.0011		
	(0.0001)	(0.0001)			(0.0009)	(0.0007)		
SB-index squared	-0.000	0.000*			-0.000	-0.000		
	(0.000)	(0.000)			(0.000)	(0.000)		
SIC-index			0.012**	0.018**			0.0039	0.013
			(0.004)	(0.007)			(0.038)	(0.044)
SIC-index squared			-0.002	-0.006*			0.025	0.037
			(0.002)	(0.003)			(0.023)	(0.025)
Capital Intensity	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Export	0.001	0.001	0.001	0.001	-0.064**	-0.061**	-0.064**	-0.059**
	(0.002)	(0.002)	(0.001)	(0.001)	(0.012)	(0.011)	(0.012)	(0.011)
Leverage (debt ratio)	-0.000	-0.000	-0.000	0.000*	-0.000**	-0.000**	-0.000**	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R&D intensity	0.146**	0.103**	0.147**	0.103**	1.122**	0.996**	-1.121**	0.992**
	(0.021)	(0.02)	(0.021)	(0.027)	(0.079)	(0.075)	(0.079)	(0.075)
Innovation intensity	0.067**	0.067**	0.067**	0.069**	0.67**	0.852**	0.67**	0.851**
	(0.01)	(0.009)	(0.017)	(0.009)	(0.059)	(0.063)	(0.059)	(0.063)
Innovation intensity sqrd	-.097**	-0.088**	-.097**	-0.089**	-0.811**	-0.989**	-0.811**	-0.988**
	(0.013)	(0.012)	(0.013)	(0.012)	(0.077)	(0.093)	(0.077)	(0.093)
Technological resources	0.016*	0.021**	0.016*	0.021**	0.236**	0.224**	0.235**	0.222**
	(0.007)	(0.006)	(0.007)	(0.006)	(0.025)	(0.034)	(0.025)	(0.034)
Labour size	0.0006	-0.006	0.001	-0.005	-0.181**	-0.23**	-0.181**	-0.23**
	(0.003)	(0.004)	(0.003)	(0.003)	(0.016)	(0.017)	(0.016)	(0.017)
Labour size squared	0.0003	0.001*	-0.000	0.0008*	0.003	0.001	0.003	0.001
	(0.0004)	(0.000)	(0.000)	(0.0004)	(0.002)	(0.002)	(0.002)	(0.001)
Economic size	-0.02**	-0.023**	-.021**	-0.022**	0.017	0.027	0.016	0.024
	(0.006)	(0.007)	(0.006)	(0.004)	(0.0215)	(0.023)	(0.021)	(0.023)
Economic size squared	0.001**	0.001**	0.001**	0.001**	0.005**	0.006**	0.005**	0.006**
	(0.000)	(0.000)	(.0003)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Firm age	0.000	0.000	0.000	0.000	-0.009**	-0.007**	-0.009**	-0.007**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.0005)	(0.0006)	(0.0005)
Average industry ROA	0.013**	0.054**	0.013**	0.05**	0.023*	0.029**	0.022*	0.029**
	(0.003)	(0.015)	(0.003)	(0.016)	(0.011)	(0.008)	(0.011)	(0.008)
Ownership type $\chi^2(5)$	43.99**	80.14**	44.96**	82.95**	17.39**	138**	17.39**	135**
Intercept	0.086**	0.105**	0.086**	0.106**	-0.004	-0.029	-0.003	-0.014
	(0.021)	(0.017)	(0.021)	(0.017)	(0.09)	(0.093)	(0.09)	(0.093)
F-value	31.46**	17.61**	31.62**	17.97**	128**	205**	128**	205**
Under-identification test[7]		$\chi^2(2)$ = 704**		$\chi^2(2)$ = 124**		$\chi^2(2)$ = 1043**		$\chi^2(2)$ = 1043**

---

[7] The Kleibergen-Paap under-identification test rejects the null hypothesis at 1% significance level, which implies the matrix is full rank and the model is identified. Thus the instruments are highly correlated with the endogenous variable.

Hansen J statistic		$\chi^2(2) =$ 4.36		$\chi^2(2) =$ 2.28		$\chi^2(2) =$ 5.78		$\chi^2(2) =$ 5.78
Observations	64478	64478	64478	64478	45213	45213	45213	45123

When the SIC-based index is used, results are consistent with our previous studies (Santarelli and Tran, 2013). Generally, highly focused firms tend to have lower profitability or, equivalently, greater diversification raises profitability. In other words, positive effects occur as firms move from a single-business strategy to a diversification strategy. However, the significant parameter of the square of the SIC relatedness index indicates the non-linear influence of diversification; the positive effects of diversification gradually fall as the firm moves further away from its core business.

With respect to the control variables, some noteworthy findings include: (i) firm profitability can be accelerated by its accumulated technological resources and R&D/innovation investments. However, the significant and negative quadratic term of the innovation investment rate indicates that profits start to fall off marginally when investments go beyond the optimal point; (ii) larger firms in terms of total assets realise lower sales than their smaller counterparts; (iii) firms in profitable industries (higher industry-level ROA) will obtain positive spillover effects from the industry to stimulate their growth and (iv), with respect to ownership types, private firms and limited liability firms outperform their state-owned and household counterparts.

As for *productivity* (TFP), estimation results are shown in the final four regressions (5 - 8) in Table 4. Nevertheless, both diversification indices do not have any significant influence on the productivities of firms. Results are consistent across the two methodological treatments. Other noteworthy findings from the productivity equation include: (i) export firms and highly indebted firms have lower productivities than their counterparts; (ii) R&D and innovation intensity significantly improve firm productivity and finally, (iii) aging and large firms are less productive.

The final robustness check limits the sample to diversified firms in manufacturing only. The regression results are reported in Appendix 5. We can observe the consistency in the findings across all treatments. In fact, the results are even a bit more statistically significant when we restrict our analysis to manufacturing firms only.

6   Conclusions

Diversification patterns are somewhat unstable and turbulent for diversified firms in Vietnam. Many firms seem to act irrationally by entering and exiting new industries in an unorganised fashion. New industries are also found to be weakly related to their own industry. In order to obtain a deeper insight into their underlying behaviours, we adopt two relatedness measures: the SB measure proposed by Teece et al. (1994) and the classical SIC-based measure to investigate (i) what actually leads firms to exit from the industry in which they recently diversified, as well as (ii) how their diversification behaviours influence their profitabilities. With respect to the methodology, we controlled for the endogeneity of diversification relatedness and the serial correlation among error terms by adopting IV-2SLS with GMM treatment.

Results suggest there has been a trend of imitation and follow-up among local decision makers in choosing the industries in which to diversify. Inexperienced firms usually resemble the direction and intensity of the diversification of dominating players within the industry (herd behaviour). Those businesses most frequently combined are not very efficient, although they generally appear more appealing to new entries. Therefore, a higher SB index does not lead to a superior entrepreneurial performance. It is however difficult to say whether exit reflects the reversal of previously inefficient decisions, or a sort of efficient form of experimentation and competence discovery (Klein et al. 2009). However, the classical SIC-based approach affirms once again our previous findings. Greater diversification raises profitability, but just to an optimum relatedness point, beyond which the positive effect starts to fade away. In conclusion, the competitive process did tend to filter out inappropriate business combinations; thus, those firms that based their diversification on the "crowd effect" will soon exit the new industry and inherently underperform in relation to their efficiency driven diversified counterparts.

Finally, we also observe the effect of experience on diversification behaviours, as identified through the increase in the statistical and economic effects of diversification relatedness on exit propensity. Although poor business development services do not equip Vietnamese firms with the necessary capabilities to choose the right industry, which induces entrepreneurs to follow the herd, they gain experience and choose more efficient business combinations over time.

References

Audretsch DB, Houweling P and Thurik AR (2000)Firm survival in the Netherlands. Review of Industrial Organization 16: 1-11.

Banerjee AV (1992) A simple model of herd behaviour. Quarterly Journal of Economics107: 797-817.

Baum FC, Schaffer ME (2003) Instrumental variables and GMM: Estimation and testing. The Stata Journal 3: 1-31.

Bernard AB, Redding SJ, and Schott PK (2011) Multiproduct firms and trade liberalization. Quarterly Journal of Economics 126: 1271-1318.

Bernardo AE, Welch I (2001) On the evolution of overconfidence and entrepreneurs. Journal of Economics and Management Strategy 10: 301-330.

Blundell R, Bond S (2000), GMM Estimation with persistent panel data: an application to production functions. Econometric Reviews 19: 321-340.

Bikhchandani S, Hirshleifer D, and Welch I (1992) A Theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy100: 992-1026.

Cassiman B, Veugelers R (2002) R&D cooperation and spillovers: Some empirical evidence from Belgium. American Economic Review 92: 1169-1184.

Chatterjee S, Wernerfelt B (1991) The Link between resources and type of diversification: Theory and evidence. Strategic Management Journal 12: 33-48.

Christensen HK, Montgomery CA (1981) Diversification strategy versus market structure. Strategic Management Journal 2: 327-343.

Clerides SK, Lach S, and Tybout JR (1998) Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico and Morocco. Quarterly Journal of Economics 113: 903-947.

Coad A, Guenther C (2014) Processes of firm growth and diversification: theory and evidence. Small Business Economics 43: 857-871.

Cohen WM, Levinthal D (1990) Absorptive capacity: A new perspective on learning and innovation. Administrative Science Quarterly35: 128-152.

Dawid H, Reimann M (2011) Diversification: a road to inefficiency in product innovations? Journal of Evolutionary Economics 21: 191-229.

DiMaggio PJ, Powell WW (1983) The iron cage revisited: Institutional isomorphism and collective rationality in organization fields. American Sociological Review 48: 147-160.

DiMaggio PJ, Powell WW (1991) Introduction. In Powell WW and DiMaggio P.J. (eds)The new institutionalism in organization analysis. University of Chicago Press, Chicago, pp. 1-38.

Edwards CD (1955) Conglomerate bigness as a source of power. Business concentration and price policy (NBER Conference Report). Princeton University Press, Princeton, NJ.

Fligstein N (1985) The spread of the multi-divisional form among large firms, 1919-1979. American Sociological Review 50: 377-391.

Gimeno J, Chen MJ (1998) The dynamics of competitive positioning: A pair-wise perspective. Academy of Management Proceedings01-08.

Greve HR, Baum JAC (2001) A multiunit, multimarket world. In Baum JAC, Greve HR (eds) Multiunit Organization and Multimarket Strategy: Advances in Strategic Management,Vol. 18, pp. 1-28. JAI Press: Oxford.

Greve HR (1998) Performance, aspirations and risky organizational change. Administrative Science Quarterly 43: 58-86.

Harhoff D, Stahl K, and Woywode M (1998) Legal form, growth and exit of West-German firms: Empirical results for manufacturing, construction, trade and service industries. Journal of Industrial Economics46: 453–488.

Haunschild PR, Miner A.S (1997) Modes of interorganizational imitations: The effects of outcome salience and uncertainty. Administrative Science Quarterly42: 472-500.

Hoskisson RE, Hitt MA, Johnson RA, and Moesel DD (1993) Construct validity of an object (entropy) categorical measure of diversification. Strategic Management Journal14: 215-235.

Hu AGZ, Jefferson GH, Jinchang Q (2005) R&D and technology transfer: Firm level evidence from Chinese industry. Review of Economics and Statistics 87: 780-786.

Jovanovic B, MacDonald GM (1994) The life cycle of a competitive industry. Journal of Political Economy 102: 322-347.

Karnani A, Wernerfelt B (1985) Multiple point competition. Strategic Management Journal 6: 87-96.

Katz ML, Shapiro C (1985) Network externalities, competition, and compatibility. American Economic Review 75: 424-440.

Klein PG, Klein SK, Lien LB (2009) Are divestitures predictable? A duration analysis. Working Paper, Contracting and Organizations Research Institute, University of Missouri.

Levinthal D (1996) Learning and Schumpeterian dynamics. In Dosi G, Malerba F (eds.)Organization and strategy in the evolution of the enterprise. MacMillan Press: London.

Lieberman MB, Asaba S (2006) Why do firms imitate each other?. Academy of Management Review 31: 366-385.

Lieberman MB, Montgomery DB (1988) First-mover advantages. Strategic Management Journal 9: 41-58.

Lien LB, Klein PG (2009a) Using competition to measure relatedness. Journal of Management 35: 1078-1107.

Lien LB, Klein PG (2009b) Diversification, industry structure and firm strategy: an organizationa economics perspective. In J.A. Nickerson and B.S. Silverman (eds), Economic Institutions of Strategy, v. 26 of Advances in Strategic Management, Emerald Group co., 289-312.

Lien LB, Klein PG (2013) Can the survivor principle survive diversification?. Organization Science 24: 1478-1494.

Loc TD, Lanjouw G, and Lensing R (2006) The impact of privatization on firm performance in a transition economy. The case of Vietnam. Economics of Transition 14: 349-389.

Markides CC (1995) Diversification, restructuring and economic performance. Strategic Management Journal16: 101-118.

Markides CC, Williamson PJ (1996) Corporate diversification and organizational structure: A resource-based view. Academy of Management Journal 39: 340-367.

Martin S. (2002) Advanced industrial economics. Blackwell Pu. Oxford.

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

Montgomery CA (1985) Product-market diversification and market power. Academy of Management Journal 28: 789–798.

Morone A (2012) A simple model of herd behaviour, a comment. Economics Letters 114: 208-211.

Nachum L (1999) Diversification strategies of developing country firms. Journal of International Management 5:115-140.

Nachum L (2004) Geographic and industrial diversification of developing country firms. Journal of Management Studies 41: 273-294.

Narduzzo A, Warglien M (1996) Learning from the experience of others: An experiment on information contagion. Industrial and Corporate Change 5: 113-126.

Ng DW (2007) A modern resource based approach to unrelated diversification. Journal of Management Studies 44: 1481-1502.

Olley GS, Pakes A (1996) The dynamics of productivity in the telecommunications equipment industry. *Econometrica*. 64: 1263-1297.

Opler TC, Titman S (1994) Financial distress and corporate performance. Journal of Finance 49: 1015-1040.

Palich LE, Cardinal LB, and Miller CC (2000) Curvilinearity in the diversification – performance linkage: An examination of over three decades of research. Strategic Management Journal21: 155-174.

Palley TI (1995) Safety in numbers: A model of managerial herd behaviour. Journal of Economic Behaviour and Organization 28: 443-450.

Pennings JM, Barkema H, and Douma S (1994) Organizational learning and diversification. Academy of Management Journal 37: 608-640.

Piergiovanni R. Carree M, and Santarelli E (2012) Creative industries, new business formation, and regional economic growth. Small Business Economics 39: 539-560.

Piscitello L (2004) Corporate diversification, coherence and economic performance. Industrial and Corporate Change 13: 757-787

Piva M, Vivarelli M (2009) The role of skills as a major driver of corporate R&D.International Journal of Manpower 30: 835 – 852.

Rumelt RP (1974) Strategy, structure, and economic performance. Harvard University Press, Cambridge, MA.

Santarelli E, Tran HT (2012) Growth of incumbent firms and entrepreneurship in Vietnam. Growth and Change 43: 638-666.

Santarelli E, Tran HT (2013) Diversification strategies and firm performance: A sample selection approach. Quaderni-Working paper DSE No. 896. Available at
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2303642

Santarelli E, Vivarelli M (2007) Entrepreneurship and the process of firm entry, survival and growth. Industrial and Corporate Change 16: 455–488.

Scharfstein DS, Stein JC (1990) Herd behaviour and investment. American Economic Review80: 465-479.

Schmalensee R (1989) Inter-industry studies of structure and performance. In Schmalensee R, Willig RD (eds.) Handbook of industrial organization.North Holland, New York.

Shepherd WG (1979) The economics of industrial organization. Prentice Hall, Englewood Cliffs, N.J.

Silverman BS (1999) Technological resources and the direction of corporate diversification: Toward an integration of the resource-based view and transaction cost economics. Management Science45: 1109-1124.

Tanriverdi H,Venkatraman N (2005) Knowledge relatedness and the performance of multibusiness firms. Strategic Management Journal 26: 97-119.

Teece DJ, Rumelt R, Dosi G, and Winter S (1994)Understanding corporate coherence: Theory and evidence. Journal of Economic Behaviour and Organization 23: 1-30.

The Economist (1997a) South Korean conglomerates. December 13, 79-80.

The Economist (1997b) China adopts the chaebol. June 7, 73-74.

The Economist (1997c) China and the chaebol. December 20, 119-120.

Tsvetkova A, Thill JC, and Strumsky D (2014) Metropolitan innovation, firm size, and business survival in a high-tech industry. Small Business Economics 43: 661-676.

Wan WP (2005) Country resource environments, firm capabilities, and corporate diversification strategies. Journal of Management Studies42: 161-182.

Weiss, L.V. (1974) The concentration-profits relationship and antitrust. In Goldschmid M, Weston S (eds) Industrial concentration: The new learning. Little Brown, Boston, MA.

Wernerfelt B (1984) A resource-based view of the firm. Strategic Management Journal5: 171-180.

Westphal J, Gulati R, and Shortell S (1997) Customization or conformity? An institutional and network perspective on the content and consequences of TQM adoption. Administrative Science Quarterly 42: 366-394.

Appendix 1: Description of Variables adopted

Indicators	Variables	Measure	Previous studies	Expected sign (with exit probability)	Expected sign (with performance)
Firm performance	ROS: Return on sales[8]	$$ROS_i = \frac{Operating\ profit_i}{Total sales_i}$$		+	
	Total factor productivity (TFP)	The semiparametric method of Olley and Pakes (1996) is used to construct the TFP variable. TFP is predicted from the production function in which sales is the production output; whereas age and capital are treated as state variables, labour and materials are freely variable inputs, and investment are taken as the proxy variable.		+	
Industry-level control variables[9]	Indusry profitability (ROA)	$ROA_i = \frac{1}{n}\sum_i^n \frac{Operating profit_i}{Total asset_i}$ (ROA_i: return on assets of industry i; n: number of firms in the industry i)	Hu et al. (2005); Santarelli and Tran (2013); Tanriverdi and Venkatraman (2005)	-/+	+
	Industry size dispersion[10]	The standard deviation of firm size from the average firm size of each four digit industry $s_n = \sqrt{\frac{1}{n}\sum_i^n (x_i - \bar{x})^2}$, where x_i is the size of firm i; $\bar{x}$: mean value of firm size of the industry; n: number of firms in the industry.	Jovanovic and MacDonald (1994), Freeman and Soete (1997)	-	
	Industry concentration[11]	The market share of the top four companies in the four digit industry $CR_4 = \sum_i^4 s_i$, CR_4: market share held by the largest 4 firms in an industry; s_i is the market share of the ith firm	Weiss (1974); Montgomery (1985); Schmalensee (1989); Lien and Klein (2009a); Lieberman and Asaba (2006)	+	
Firm-level control variables	R&D intensity/ Innovation investment rate[12]	$R\&D\ intensity_i = \frac{R\&D\ expenditure_i}{business\ expenditure_i}$; $Innov.rate_i = \frac{Investments\ in\ innovation_i}{Total\ annual\ investment_i}$	Cohen and Levinthal (1990); Levinthal (1996); Santarelli and Tran (2013)	-	+

[8]The rationales for using ROS rather than the widely used logarithm of profit and ROA include: the logarithm of profit excludes firms operating at a loss (negative profit) from the analysis and assets would carry book values and require a longer time of availability. For growth measures, an attempt was made to use the growth of sales as a dependent variable to reflect a firm's performance, but the test for serial autocorrelation indicated a significant violation in the assumption of no serial correlation, which seriously biased the final findings with the panel data.

[9]Industry-level characteristics are aggregated at the four-digit ISIC level. Industry variables are constructed for both the main industry and the newly-entered industry.

[10] Industry size dispersion is proxied for the effect of industry life cycle. A small standard deviation indicates a technological convergence, which is typical of mature industries; on the contrary, a large standard deviation suggests it is possible to adopt a wide range of technologies and market strategies, as in the case of the growing stages of an industry

[11]We use an interaction term between industry concentration and relatedness measures as a test for rivalry-led imitation (mutual forbearance) according to what is suggested by Lien and Klein (2009a) and Lieberman and Asaba (2006)

[12]We allow for a non-linear relationship by including a squared innovation investment rate in the regression

	Technological resources	$Tech.\,resources_i = \dfrac{No.\,of\,technical\,employees_i}{Total\,labor\,force_i}$	Cohen and Levinthal (1990); Cassiman and Veugelers (2002); Piva and Vivarelli (2009); Santarelli and Tran (2013)	-	+
	Firm size[13]	Labour size: natural logarithm of the number of total employees; Economic size: natural logarithm of total assets	Palich et al. (2000); Santarelli and Tran (2013)	-	-/+
	Firm age	Number of years firm has been operating	Piergiovanni et al. (2012); Santarelli and Tran (2013); Tsvetkova et al. (2014)	-/+	-/+
	Firm export	The dummy attains 1 if the firm exports, 0 otherwise	Clerides et al. (1998); Bernard et al. (2011); Santarelli and Tran (2013)	-/+	+
	Debt ratio	$Debt\,ratio_i = \dfrac{Total\,debt_i}{Total\,assets_i}$	Opler and Titman (1994); Santarelli and Tran (2013)	-	-
	Capital intensity	$Capital\,intensity_i = \dfrac{Fixed\,assets_i}{Number\,of\,employees_i}$	Shepherd (1979); Audretsch et al. (2000); Santarelli and Tran (2013)	+	-
	Ownership types	Six dummies is to control for state-owned firms, partnerships and cooperatives, private firms, limited liability firms, joint stock firms	Harhoff et al. (1998); Santarelli and Tran (2013)		
	Productivity	The interaction term between TFP from Olley and Pakes (1996) model and relatedness measures	Olley and Pakes (1996)	-	
Relatedness variables	Survivor-based index[14]	$SB_i = \dfrac{\sum SR_{ij}s_j}{\sum s_j}$ (SR_{ij}: standardised measure of how much the actual number of combinations exceeds the random combination; s_j: sales from industry j)	Teece et al. (1994); Lien and Klein (2009a); Lien and Klein (2013)	-	+
	SIC-based index[15]	$SIC_i = \dfrac{\sum d_{ij}s_j}{\sum s_j}$, where $d_{ij} = 2$ if i and j are in the same three-digit SIC code, $d_{ij} = 1$ if i and j are in different three-digit but in the same two-digit SIC codes and $d_{ij} = 0$ if i and j are in different two-digit SIC codes; s_j: sales from industry j	Hoskisson et al. (1993); Markides and Williamson (1996), Palich et al. (2000); Santarelli and Tran (2013).	-	+

[13] A quadratic term is also added to establish a non-linear relationship between a firm's size and its diversification performance

[14] The detailed construction is in Appendix 3.

[15] This relatedness measure is used as a good benchmark for SB measure, as it shows the relatedness in terms of industrial codes rather than the frequency of industry combinations; hence, it does not reflect herding behaviour.

Appendix 2: Review of diversification measures

Type of measures	Measures	Formula	Advantages	Disadvantages	References
Categorical approach	Rs/Rr: the proportion of a firm's revenues attributable to its largest single business/group of related business	4 major categories: single business ($Rs \geq 0.95$ & $Rr \geq 0.7$), dominant business ($0.7 \leq Rs < 0.95$ & $Rr \geq 0.7$), related business ($Rs < 0.7$ & $Rr \geq 0.7$), and unrelated business ($Rs < 0.7$ & $Rr < 0.7$)	Simple	- Subjectivity - Requires detailed business-level information => time consuming - relies heavily on qualitative assessment of diversification patterns	Wrigley (1970); Rumelt (1974)
	Two-dimensional count measure	Narrow spectrum diversification (NSD): number of 4-digit SIC codes that a firm participates Broad-spectrum diversification (BSD): number of 2-digit SIC codes that a firm participates	Simple	- Subjectivity - Reliability problems - Time consuming	Varadarajan and Ramanujam (1987); Hall and John (1994)
Count approach	Herfindahl Concentric measure	$\sum_j m_{ij}^2$, m_{ij}: proportion of jth classified group to ith firm's total sales	Objectivity, able to deal with degree of diversity	- Reflects concentration rather than diversification	Montgomery (1982); Chatterjee and Blocher (1992)
	Berry-Herfindahl index	$$1 - \left(\sum_j m_{ij}^2\right) / \left(\sum_j m_{ij}^2\right)^2$$ m_{ij}: proportion of jth classified group to ith firm's total sales	Objectivity, able to deal with degree of diversity	- Takes in to account only the variation in turnover.	Caves et al. (1980)
	Entropy index	$$\sum_j m_{ij} \, ln(1/m_{ij})$$ m_{ij}: proportion of jth classified group to ith firm's total sales	Objectivity, able to deal with degree of diversity	- Strong dependence on the significance of activity areas Content validity is questionable (Robins and Wiersema, 2003)	Jacquemin and Berry (1979); Kim (1989); Palepu (1985)
Recent developments	Skill-based relatedness	Similarities among occupational profiles of different industries	Objectivity and reliability	Just capture a dimension of relatedness Focus only on specific industries	Farjoun (1994); Neffke and Henning (2012)
	Technological relatedness	Similarity of technological resources, for ex. patents, of different industries; or labour sharing across industries	Objectivity and reliability	Just capture a dimension of relatedness Focus only on specific industries	Robins and Wiersema (1995)
	Survivor-based relatedness	How often industries are actually combined to what one would expect if diversification patterns were random	Directly reflects diversification behaviours; Higher predictive validity	Fails to capture relatedness if firms act boundedly rational	Teece et al. (1994); Lien and Klein (2009a)

Appendix 3 – The survivor-based (SB) index of relatedness

Let the population of the diversified firms consist of K firms, each active in two or more industries i. Let $Z_{ik} = 1$ if a firm n is active in the industry i. The number of industries participated in by a firm k is $m_k = \sum_i Z_{ik}$, and the number of diversified firms present in the industry i is $n_i = \sum_k Z_{ik}$. Let C_{ij} be the number of diversified firms active in both industries i and j, such that $C_{ij} = \sum_k Z_{ik}Z_{jk}$. In other words, C_{ij} is a count of how often industries i and j are actually combined within the same firm. C_{ij} will be larger if industries i and j are related, but it will also increase with n_i and n_j. To avoid the effect of the sizes of the industries i and j, the number C_{ij} is compared with the number of expected combinations if diversification patterns were random. The random diversification hypothesis can be represented as a hypergeometric distribution function $\Pr(X_{ij} = x) = \dfrac{\binom{n_i}{x}\binom{K-n_i}{n_j-x}}{\binom{K}{n_j}}$, where x is the number of firms active in both industries i and j, and n_i and n_j are drawn independently and randomly from a population of K firms. The mean and variance of X_{ij} are, respectively, $\mu_{ij} = E(X_{ij}) = \dfrac{n_i n_j}{K}$; $\sigma_{ij}^2 = \mu_{ij}\left(1 - \dfrac{n_i}{K}\right)\left(\dfrac{K}{K-1}\right)$. Then, the weighted average SB relatedness of the target industry i to all other industries in the firm is then defined as $SB_i = \dfrac{\sum SR_{ij}s_j}{\sum s_j}$, wherein $SR_{ij} = \dfrac{C_{ij}-\mu_{ij}}{\sigma_{ij}}$ is a standardised measure of how much the actual number of combinations exceeds the expected combinations under the random diversification hypothesis, and s_j is the sales of a diversified firm in industry j.

Appendix 4: Means, Standard Deviations, and Correlation Coefficients of Independent Variables

	Mean	Std. Dev	Indus MES	Indus ROA	Indus Conce	Capital intensit	Export	Labour size	Eco. Size	Debt ratio	RD intensity	Innov invest	Tech resou	Firm age	SIC related	SB related
Industry var.																
Indus. MES	82.48	208.2	1.000													
Indus. ROA	0.027	0.373	0.042*	1.000												
Indus. Concentration	0.28	0.196	0.174*	0.011*	1.000											
Firm variables																
Capital intensity	119.3	1132.5	0.003	0.016*	0.024*	1.000										
Export	0.148	0.356	0.059*	-0.006	0.011*	0.004	1.000									
Labour size	3.443	1.551	0.260*	0.011*	0.077*	-0.014*	0.218*	1.000								
Economic size	8.228	1.905	0.176*	0.012*	0.053*	0.113*	0.232*	0.778*	1.000							
Debt ratio	2.855	320.5	0.000	-0.000	0.003	0.031*	0.011*	0.003	0.013*	1.000						
R&D intensity	0.012	0.079	0.037*	0.001	0.037*	0.003	-0.019	0.044*	0.039*	0.001	1.000					
Innovation investment	0.121	0.243	-.018*	-0.003	-0.04*	-0.01*	-0.08*	-.073*	-.088*	-0.003	0.052*	1.000				
Technological resources	0.071	0.174	-.016*	-.012*	0.028*	0.028*	-.046*	-.039*	0.042*	-0.001	-0.026*	0.01*	1.000			
Firm age	7.518	8.179	0.094*	0.011*	0.081*	-0.005	0.124*	0.327*	0.288*	-0.001	0.012*	-.157*	-.001	1.000		
Relatedness variables																
SIC-based relatedness	0.067	0.3	-.023*	0.006	0.025*	0.004	-.011*	0.001	0.002	-0.001	0.001	-.025*	-.018*	-0.000	1.000	
SB-based relatedness	1.964	11.126	0.034*	-0.005	0.175*	0.014*	0.007	0.098*	0.083*	-0.000	0.021*	-.043*	-.003	0.056*	.185*	1.000

Note: *: significant at the 1% level. Observations: 67522.

Appendix 5: Robustness check - Profitability of diversified manufacturing firms

	Diversified Firms' Profitability (ROS)			
	SB-based relatedness index		SIC-based relatedness index	
	GLS[1]	IV-GMM[2]	GLS[1]	IV-GMM[2]
SB-index	0.0001(0.0001)	0.0001(0.0001)		
SB-index squared	-0.000(0.000)	0.000*(0.000)		
SIC-index			0.037*(0.006)	0.005* (0.003)
SIC-index squared			-0.003 (0.004)	-0.003 (0.003)
Capital Intensity	-0.000** (0.000)	-0.000(0.000)	-0.000** (0.000)	-0.000(0.000)
Export	0.0004 (0.002)	0.001(0.002)	0.000 (0.002)	0.001 (0.002)
Leverage (debt ratio)	-0.000** (0.000)	-0.000** (0.000)	-0.000** (0.000)	0.000** (0.000)
R&D intensity	0.149** (0.011)	0.143**(0.025)	0.149** (0.011)	0.144** (0.025)
Innovation intensity	0.065**(0.012)	0.071**(0.013)	0.065** (0.012)	0.071** (0.014)
Innov intensity sqrd	-0.075**(0.014)	-0.082**(0.016)	-0.075** (0.014)	-0.082** (0.017)
Technological resources	0.03**(0.007)	0.035**(0.013)	0.03**(0.007)	0.058** (0.02)
Labour size	0.011** (0.003)	0.012** (0.005)	0.011** (0.003)	0.012** (0.005)
Labour size squared	-0.0007* (0.0004)	0.001*(0.000)	-0.001* (0.000)	-0.001* (0.0005)
Economic size	-0.044**(0.004)	-0.046** (0.005)	-0.044** (0.004)	-0.046** (0.005)
Economic size squared	0.002**(0.0002)	0.002**(0.0003)	0.002** (0.0002)	0.002**(0.000)
Firm age	0.000(0.000)	0.000(0.000)	0.000(0.000)	0.000(0.000)
Average industry ROA	0.007**(0.002)	0.008**(0.001)	0.007** (0.002)	0.008** (0.001)
Ownership type $\chi^2(5)$	58.61**	102.26**	58.67**	82.95**
Intercept	0.049**(0.018)	0.058** (0.02)	0.05** (0.018)	0.058** (0.02)
F-value	38.29**	24.17**	38.27**	24.17**
Hansen J statistic		$\chi^2(2) = 0.005$		$\chi^2(2) = 0.025$
Observations	29228	29228	29228	29228

Notes: 1: Prais-Winsten and Cochrane-Orcutt generalised least squares correcting for first-order serial correlation and heteroskedasticity with robust standard errors. 2: IV-2SLS with GMM treatment, clustering across firms. 3: We control for provincial fixed effects by three dummies *south, central* and *north* for isolating those provinces located in the North, Centre and South of Vietnam, respectively.

*: significant at the 5% level; **: significant at the 1% level. Standard errors in parentheses.