Does it really matter how a firm diversifies? Assets-in-place diversification versus options-based diversification

Pablo de Andrés [a], Gabriel de la Fuente [b] and Pilar Velasco [c*]

(This version: 10th March 2016)

[a] Universidad Autónoma de Madrid, Department of Finance, Faculty of Economics, Cantoblanco Campus, c/Adam Smith 2 - 28049 Madrid (Spain). Phone: +34 91 497 6527. E-mail: p.andres@uam.es

[b] Universidad de Valladolid, Department of Financial Economics, Faculty of Economics, Avda. Valle del Esgueva 6 – 47011 Valladolid (Spain). Phone: +34 983 423334. E-mail: gabriel.fuente@uva.es

[c*] **Corresponding author**. Universidad de Alcalá, Department of Economics and Business, Faculty of Economics, Plaza Victoria 2 – 28802 Alcalá de Henares, Madrid (Spain). Phone: +91 885 52 17. E-mail: mpilar.velasco@uah.es

Does it really matter how a firm diversifies? Assets-in-place diversification versus options-based diversification

(This version: 10th March 2016)

Abstract

This paper analyzes whether the valuation effect of corporate diversification depends on how this strategy is implemented. According to the real options approach, two extreme diversification patterns may be identified: one based on full exercise of available options (an assets-in-place diversification), and another aimed at seeding multiple growth options in subsequent businesses (options-based diversification). We propose an index to approximate the distance to these two diversification patterns and we explore its impact on firms' value for a sample of U.S. firms during 1998-2010. We find that as a firm's diversification approaches an options-based pattern, it becomes a more value-enhancing strategy.

1. Introduction

The diversification-value linkage has constituted a prolific area for research. Existing literature offers abundant yet inconclusive evidence, with the research question having become widely known as the *diversification puzzle*[1]. Even though the bulk of the research provides evidence that diversified firms trade at a discount relative to non-diversified companies in their industries (Lang and Stulz, 1994; Berger and Ofek, 1995; Servaes, 1996; Stowe and Xing, 2006; Borghesi, Houston and Naranjo, 2007; Hoechle *et al.*, 2012), other works call these findings into question, and report a non-statistically significant relationship (Villalonga, 2004b; Elsas, Hackethal and Holzhäuser, 2010), a quadratic relationship (Palich, Cardinal and Miller, 2000), or even premiums for diversifying (Campa and Kedia, 2002; Villalonga, 2004a).

Recently, research has suggested that these inconsistent findings might be due to three reasons: biases in data (Villalonga, 2004a), methodological issues (Campa and Kedia, 2002; Villalonga, 2004b), and the presence of moderating factors in the diversification-value relationship. Moderating factors suggest that the impact of this strategy on performance may not be homogeneous across firms, but dependent on aspects that might enable certain enterprises to create more value than others. Such moderating variables could be classified into three categories: market and institutional level factors (Fauver, Houston and Naranjo, 2003; Rudolph and Schwetzler, 2013), industry level factors (Santaló and Becerra, 2008), and firm level factors (Rajan, Servaes and Zingales, 2000).

Among the firm-specific characteristics that may account for differences in diversification value outcomes, a strand of literature explores the role of a firm's growth opportunities. Bernardo and Chowdhry (2002) attribute a significant role to growth

---

[1] See Martin and Sayrak (2003) for a survey.

opportunities when accounting for diversification discount: unisegment firms have more options to expand whereas diversified firms may have exhausted part of these. Concurring with this line of argument, Borghesi *et al.* (2007) claim that pure-play firms and their diversified industry-peers exhibit different growth potential. Once the age of the firm, used as a proxy for growth opportunities, is controlled for, the initially displayed discount decreases. Ferris *et al.* (2002) analyse diversification for a sample of international joint ventures and show that diversification is only value-destroying in enterprises that have a poor set of growth opportunities. Holder and Zhao (2015) claim that prior evidence on diversification discount neglects the real options impact on the value measures. They find that diversification discount may be the joint result of the value increase in below average performers arising from having explored new opportunities through unrelated diversification, and the value decrease in above average performers from exploiting excess capability through the same strategy. On the other hand, Stowe and Xing (2006) offer evidence that differences in growth opportunities are not the main driver of the diversification discount, since the discount persists even after such differences are controlled.

In light of these findings, it seems worth examining in greater depth the role of growth options in the diversification-value puzzle. Specifically, the questions which arise are: firstly, whether diversification strategies can differ depending on their effect on shaping a firm's RO portfolio; and secondly, whether such differences might have an impact on corporate value. We adopt a RO perspective to identify two extreme diversification patterns: one based on fully-committed entry into new markets (assets-in-place diversification), and another aimed at seeding multiple growth options in subsequent businesses (options-based diversification). These two investment patterns translate to opposing ways of diversifying: assets-in-place diversification (AiPD) versus option-based diversification (OD), respectively. Logically, no

firm adopts either of these 'extreme' diversification patterns but rather an intermediate one which might be closer to one or the other. Measuring the proximity to one of these extreme patterns is no easy task, since neither a firm's strategy nor its RO portfolio are directly observable.[2] To address the problem of unobservability, we propose a proxy which can reflect some of the key effects to emerge from each strategy. To the best of our knowledge, no prior research has attempted to proxy this kind of diversification pattern[3]. Taking the concept of fit as *profile deviation* (Venkatraman, 1989), we propose a two-dimensional index which reflects the degree of diversification asymmetry across a firm's different businesses through an inter-segment measure of investment distribution and an intra-segment measure based on industry peers in each sector.

The question which arises afterwards is whether the proximity to one of these extremes makes any difference in the value of diversification. Amihud and Lev (1981) argue that the critical issue on the valuation effect of diversification is what kind of risk is reduced by diversification and whether stockholders can diversify it in their individual portfolios. Were investors able to diversify at a lower cost than enterprises, corporate diversification would destroy value. However, diversification aimed at providing the firm with growth options might not be so easily replicated, thus resulting in a diversification premium. A growth options diversification boosts flexibility to adjust decisions as uncertainty is resolved, and is geared towards not only exploiting but also exploring and generating further opportunities in new industries before fully committing. As a result of this flexibility, corporate diversification may reduce risk and serve as a 'strategic insurance' (Raynor, 2002) which cannot be replicated by investors. The most an individual investor could hope to achieve is to

---

[2] As is well known, unobservability is by no means an unusual issue in Social Sciences, since most variables in their models are unobservable and proxies are used instead.

[3] The only related precedent is Klingebiel and Adner (2015) who develop a classification of product innovation strategies to distinguish between the real options logic and other alternative resource allocation regimes.

replicate the growth options portfolio by acquiring those stocks which contain said options. However, the value of this replicated portfolio should be less than the value of the growth options portfolio of the diversified firm, since optimal joint exercise of an options portfolio always proves more efficient than the sum of the individual optimal exercise of each option. Such arguments lead us to hypothesize that RO-oriented diversification might be a value-enhancing pattern of diversification.

Some results in prior research may support this supposed superiority of the OD pattern for creating value over AiPD. For example, Teplensky *et al.* (1993) find that incremental strategies lead to better performance in uncertain and dynamic environments such as emerging markets since they avoid full commitment of resources while past performance acts as a feedback mechanism for future strategic decisions. In a similar vein, Andreou, Louca and Petrou (2016) report a discount in enterprises moving one-time from a single segment to multiple ones, as opposed to a premium in diversifiers which undertake this strategy several times. Mitton and Vorkink (2010) also concur with these findings showing that the valuation effect of diversification is positively related to return skewness, which is consistent with superiority of the OD pattern insofar as positive skewness of a firm's stock returns approximates RO relevance in its total assets mix (Trigeorgis and Lambertides, 2014).[4] Overall, this empirical evidence suggests that OD pattern may be more value-enhancing than AiPD.

Based on this logic, we assess whether these diversification patterns explain the discounts/premiums associated to this strategy. Our analysis is conducted on an unbalanced panel sample of 3,165 U.S. firms (16,554 firm-year observations) for the 1998-2010 period.

---

[4] Mitton and Vorkink's explanation is not based on RO. They argue that corporate diversification reduces stock return skewness as a consequence of returns' compensation, similar to what happens when combining stocks in a portfolio. Should investors prefer positive skewness, a firm's relative value would be discounted as skewness is reduced by segment diversification.

Empirical models are estimated using the Heckman two-step procedure to control for self-selection of the diversification decision. We find strong evidence concerning the relevance of the diversification pattern to explain diversification value outcomes. Our empirical findings clearly show that options-based diversification emerges as a more value-creating strategy.

The rest of the paper is structured as follows. Section 2 explains our proposed diversification patterns and develops a proxy to measure them. The following section describes our empirical models, estimation methodology, data set, and variables. Section 4 and 5 describe our main empirical results and robustness tests. To round off the paper, a discussion of the implications of the findings, limitations, and directions for future research is offered.

2. RO diversification patterns and proxies

Traditional diversification indexes are geared towards capturing the scope of diversification in terms of distributing firms' business activity across their segments. Yet, by themselves they fail to provide information as to the way of investing. Degree and pattern of diversification constitute two different features of this strategy, each requiring specific measures.

From the Real Options (RO) lens, two contrasting investment paths can be distinguished (Bowman, Hurry and Miller, 1992): a *one-step strategy*, which mainly involves full-scale commitments by making large sunk investments, versus a *growth option strategy*, entailing minor commitments in strategic areas which serve as platforms for future investments. These two investment patterns translate to opposing ways of diversifying: assets-in-place diversification (AiPD) versus option-based diversification (OD), respectively.

In AiPD, the firm holds a large participation in each business it is involved in. Diversification into a new business is conceived as a one-shot investment strategy, meaning immediate exercise or abandonment of previously acquired growth options. This strategy allows firms fully-committed entry into new markets to exploit potential economies of scope and synergies, at the expense of taking a greater risk in each commitment, losing the flexibility to readjust the strategy along the way, and achieving limited exploratory capacity development. This diversification path prioritises exploiting available opportunities rather than keeping them open to wait for the best moment to exercise. AiPD may correspond to a greater extent to the traditional notion of diversification under which each diversification movement "consumes" a firm's growth options in return for achieving strategic advantages such as synergies and market power.

In OD, the main objective of diversification is to develop further strategic options in new businesses (Williamson, 2001). Each investment is regarded as "*a foothold in preparation for the next decision*" (Bowman *et al*., 1992: 98). The firm undertakes small-scale entries into several businesses, which is seen as acquiring an option that can act as a 'platform' for future growth opportunities (Kogut and Kulatilaka, 1994). As investment conditions evolve, firms will maintain, expand, exercise, or abandon these options, while acquiring new options to diversify and keeping them open in other areas. As a result, OD implies more active and flexible management of the strategy to react to uncertainty and fresh information, as well as enabling firms to gain experience in a new field and explore further opportunities before fully committing. This "wait and see" logic is by no means free of costs such as risk of pre-emption or loss of first-mover advantages due to the undeveloped participation in the new businesses, which delays any major commitment. However, OD allows firms to continuously build and maintain a portfolio of strategic options for the future (Williamson, 2001), thereby

encouraging experimentation and learning. Overall, OD means simultaneously exploring and exploiting growth options, which is likely to enhance the value of the firm's growth options portfolio.

In line with these two diversification patterns identified from an RO perspective and the key effects to emerge from them, we propose an index to proxy for. Our proposal incorporates two dimensions: an inter-segment dimension (*INTER*) and an intra-segment dimension (*INTRA*).

2.1. Inter-segment dimension (INTER)

The *INTER* measures the degree of inequality in the distribution of the firm's level of diversification across its different businesses. An AiPD will translate to a more uniform distribution of the firm's activity (measured in terms of sales) across the different segments since it holds a more balanced commitment in the various industries. In contrast, an OD will reflect unequal distribution, mostly concentrating its participation in core businesses coupled with minor exploratory investments in new industries. Overall, *INTER* is intended to offer such an overview of the investment strategy followed by the firm in its overall businesses portfolio. We approximate this dimension by a Gini index, computed as follows[5]:

$$
INTER = \frac{\sum_{s=1}^{n-1}(p_s - q_s)}{\sum_{s=1}^{n-1}p_s} \tag{1}
$$

where *s* stands for each firm's segment (s ranges from 1 to n), *n* represents the total number of firm segments, p_s denotes the cumulative proportion of sales (from segment 1 to segment *s*), q_s denotes the cumulative proportion of total diversification (from segment 1 to segment *s*) calculated as:

---

[5] To compute the Gini index, each firm's segments first needs to be sorted into ascending order of sales.

9

$$q_s = \sum_{h=1}^{s} \frac{P_h^2}{HERF} \qquad (2)$$

where P_h denotes the proportion of sales in each segment h and $HERF$ is the Herfindahl index

defined by Hirschman (1964): $\sum_{h=1}^{n} P_h^2$

INTER takes values between 0 and 1. An *INTER* equal to zero reflects perfect equality,

and the higher the *INTER*, the greater the inequality. Thus, the nearer the *INTER* is to zero,

the closer the firm's diversification pattern is to AiPD; whereas the nearer the *INTER* is to

one, the closer the firm's diversification pattern is to OD.[6]

2.2. Intra-segment dimension (INTRA)

To appropriately evaluate whether a firm's diversification profile is closer to either AiPD

or to OD, we need to scale its participation in each business. With this aim in mind, we

incorporate an industry comparative framework based on the average of its industry peers in

that sector. We use a multiplier approach to estimate the sales each firm would obtain from

each business segment (imputed sales) were it to follow average industry commitment. We

adopt a similar procedure to Berger and Ofek's (1995) methodology for assessing a firm's

imputed value. First, we take all public listed firms (both single-segment and diversified)

operating in each industry j, and calculate each firm i's ratio of i's sales in industry j over

firm i's total assets. Sales are scaled by total assets to make the commitment of different sized

enterprises comparable. Next, we compute the mean ratio for each industry j at the 4-digit

code level. We then multiply this industry mean multiple by the firm's total assets to

calculate imputed sales for each segment s of each firm i.

$$imputed\_segment\_sales_{is} = TA_i * \overline{multiple}_s(S/TA) \qquad (3)$$

---

[6] For single-segment firms, we assume the *INTER* to equal zero.

where TA_i represents firm i's total assets, and $\overline{multiple_s(S/TA)}$ denotes the mean multiple of the corresponding industry j of segment s.

To evaluate the firm's scale of participation relative to the average of its industry peers, we compute a *commitment ratio* which compares firm $i's$ real sales figures in each segment s against its corresponding imputed sales:

$$commitment\_ratio_{is} = \frac{firm's\_segment\_sales_{is}}{firm'\_imputed\_segment\_sales_{is}} \qquad (4)$$

A commitment ratio above or equal to 1 will indicate that firm i holds an above-average commitment in the industry and thus pursues an AiP pattern of investment in that sector. Otherwise, it will display an option pattern based on under-developed participation in that business, seen as the acquisition of an option which may serve as a platform for further opportunities.

Next, we compute the intrasegment component (*INTRA*) of our index, which seeks to capture the company's overall degree of commitment in all its business segments. *INTRA* is measured by the ratio of the number of firm i's segments displaying commitment ratios above or equal to 1 over firm i's total number of segments. *INTRA* is positively related to the AiP pattern of diversification: the closer the firm is to AiPD, the higher the *INTRA,* the latter's value ranging between 0 and 1. If the company followed an AiP strategy in all its businesses, *INTRA* would by definition be 1.

2.3. The diversification pattern index (DIVPAT)

Finally, we combine *INTRA* and *INTER* in a two-dimensional index capturing the diversification pattern (*DIVPAT*), which is devised on the basis of the Euclidean distance. Following Venkatraman (1989: 433), we take a concept of *fit* as profile deviation to analyse

the degree of adherence of a firm's diversification pattern to an externally specified profile. In our analysis, that externally specified profile serving as a reference will be the extreme case of AiPD, and the index to capture RO patterns based on the Euclidean distance can generally be expressed as:

$$DIVPAT = \sqrt{(INTER_i - AiP\_INTER)^2 + (INTRA_i - AiP\_INTRA)^2} \qquad (5)$$

where $INTER_i$ denotes the value of firm i's *INTER* variable, AiP_INTER the value of the *INTER* for the extreme case of AiPD, $INTRA_i$ the value of firm i's *INTRA* variable, and AiP_INTRA the value of the *INTRA* for the extreme case of AiPD.

The reference profile can be specified either theoretically or empirically (Venkatraman, 1989). From a theoretical standpoint, in an extreme AiPD, *INTER* would equal 0, representing perfect equality, whereas *INTRA* would equal 1, indicating that every firm's segment holds a commitment above the industry mean. Thus, our diversification pattern index is defined as the deviation (Euclidean distance) of the firm's diversification pattern (in both its *INTER* and *INTRA*) from the extreme AiPD (*INTER*=0, *INTRA*=1):

$$DIVPAT(THEO) = \sqrt{(INTER_i - 0)^2 + (INTRA_i - 1)^2} \qquad (6)$$

As *DIVPAT(THEO)* represents the degree of deviation of the firm's diversification path from the case of extreme AiPD (theoretically specified), the higher this index, the closer a firm's diversification profile is to the OD pattern; while the lower this index, the closer to the AiPD pattern.

3. Research design: Data, variables and econometric models

3.1. Database, sample selection and description

12

Our initial sample comprises all publicly traded US companies (16,637 firms) included in the Worldscope database during the period 1998[7]-2010. To mitigate survivorship bias problems, we consider both active and currently inactive firms[8]. We use Worldscope as our premier source of data, and employ annual data at both the 4-digit SIC code industry segment and company level. Two additional information sources are used to supplement this data: Datastream, to collect market data, and the Bureau of Economic Analysis (U.S. Department of Commerce), to obtain macroeconomic data.[9]

As shown in Table 1, we select the sample by applying Berger and Ofek's (1995) criteria to ensure that our results are comparable to prior literature. First, we drop firm-years with any segment operating in the financial services industry (SIC codes 6000-6999) or with non-positive sales. In addition, we remove observations with missing data on total capital, total sales, and segment-level sales. We also require firms to have total sales greater than or equal to $20 million, and the sum of segment sales to be within the range of 99% to 101% of a firm's reported total sales. All these Berger and Ofek (1995) requirements restrict our dataset to 28,206 firm-year observations, 67% corresponding to focused firms and 33% to diversifiers. Finally, we remove outlying observations of the main variables included in our analysis. Our final dataset for estimation purposes is an unbalanced panel sample of 16,554 firm-year observations, comprising a total of 3,165 companies for the 1998-2010 period.

TABLE 1 ABOUT HERE

---

[7] As of December 15, 1997, the new SFAS 131 reporting standard became effective for fiscal years in the United States, replacing the previous SFAS 14. Our sample starts in 1998 to ensure homogeneity of data.

[8] Firms which disappeared from the sample during the analysis period for various reasons such as mergers or bankruptcy.

[9] Bureau of Economic Analysis data is taken from official website: http://www.bea.gov/national/index.htm. Worldscope and Datastream data is obtained from ThomsonOne package by Thomson Reuters.

Table 2 presents full-period general descriptive statistics concerning the financial profile of the companies in the final sample. As can be seen, there is substantial heterogeneity across firms in certain characteristics such as size (either approximated by total sales, total assets, or market capitalisation), performance (measured by EBIT), and debt.

TABLE 2 ABOUT HERE

3.2. Econometric approach, empirical models and variables

Our estimation methodology is the Heckman two-step procedure to control for self-selection (Heckman, 1979). Selectivity appears when diversification is not assigned randomly across companies, with firms either self-selecting to diversify or to remain focused (Campa and Kedia, 2002; Villalonga 2004b; Miller, 2006). Factors affecting firms' propensity to diversify may also impact diversification value outcomes. If so, diversification variables would be correlated with the error term in the diversification-value models, and OLS estimators would not prove consistent. The Heckman two-stage method considers this self-selection bias as an omitted variable problem and corrects for it.

The first step of the Heckman estimation involves a probit analysis to model the firm's propensity to diversify (*selection equation*). It enables us to obtain self-selection correction in the form of the inverse of Mill's ratio (λ) (Greene, 2003), which will be included at the second stage to correct for selection bias. The resulting estimators of this latter equation would thus reflect the net effect of the diversification strategy on the dependent variable once sample selectivity has been corrected.

Following Campa and Kedia (2002), we consider the following selection equation:

$D_{it} = \gamma_0 + \gamma_1 \, LTA_{it} + \gamma_2 \, EBITsales_{it} + \gamma_3 \, CAPEXsales_{it} + \gamma_4 \, PNDIV_{it} + \gamma_5 \, PSDIV_{it}$

$$+\gamma_6 \text{ChangeGDP}_{it} + \gamma_7 \text{CONTRACTION}_{it} + \eta_{it} \tag{7}$$

being $D_{it}=1$ if $D_{it}^*>0$, and $D_{it}=0$ if $D_{it}^*<0$, where D_{it}^* is an unobserved latent variable seen to be $D_{it}=1$ if $D_{it}^*>0$ (diversified firm), and equalling zero otherwise (unisegment firm), and η_{it} is an error term. The dependent variable in our selection equation is a diversification dummy which equals 1 if the firm operates in two or more different 4-digit SIC industries, and zero otherwise. To ensure comparability with prior research, we assume the diversification decision to be driven by the following characteristics[10]:

- at firm-level: firm size, estimated by the natural logarithm of the book value of total assets (*LTA*); profitability, approximated by the ratio EBIT to sales (*EBITsales*); and the firm's level of investment in current operations, proxied by the capital expenditures to total sales ratio (*CAPEXsales*).

- at industry-level: industry attractiveness, based on both the fraction of firms in the firm's core industry that are diversified (*PNDIV*) and the proportion of the firm's core industry sales accounted for by diversifiers (*PSDIV*)[11].

- and at the macro-economic level: economic cycle attractiveness, approximated by the real growth rates of gross domestic product, calculated as the GDP percent change based on 2005 dollars (*changeGDP*); and the number of months in the year the U.S. economy was in recession (*CONTRACTION*).

At the second stage of the Heckman procedure, our main models (*outcome equations*) are estimated by ordinary least squares (OLS)[12]. First, as preliminary analyses to test the validity of our index, we relate both *INTER* and *INTRA* to a firm's value (equations (8) to (10)):

---

[10] See Campa and Kedia (2002) for a further explanation of the variables selection.
[11] We calculate these two proxies at the 4-digit SIC level.

$$\text{EXCESS\_VALUE}_{it} = \alpha + \beta_1 \text{INTER}_{it} + \beta_2 \text{LTA}_{it} + \beta_3 \text{LDTA}_{it} + \beta_4 \text{EBITsales}_{it}$$

$$+ \beta_5 \text{CAPEXsales}_{it} + \beta_6 \text{LTA2}_{it} + \beta_7 \lambda_{it} + \beta_8 \text{Industry\_dummies}_{it}$$

$$+ \beta_9 \text{Year\_dummies}_{it} + v_{it} \tag{8}$$

$$\text{EXCESS\_VALUE}_{it} = \alpha + \beta_1 \text{INTRA}_{it} + \beta_2 \text{LTA}_{it} + \beta_3 \text{LDTA}_{it} + \beta_4 \text{EBITsales}_{it}$$

$$+ \beta_5 \text{CAPEXsales}_{it} + \beta_6 \text{LTA2}_{it} + \beta_7 \lambda_{it} + \beta_8 \text{Industry\_dummies}_{it}$$

$$+ \beta_9 \text{Year\_dummies}_{it} + v_{it} \tag{9}$$

$$\text{EXCESS\_VALUE}_{it} = \alpha + \beta_1 \text{INTER}_{it} + \beta_2 \text{INTRA}_{it} + \beta_3 \text{LTA}_{it} + \beta_4 \text{LDTA}_{it} + \beta_5 \text{EBITsales}_{it}$$

$$+ \beta_6 \text{CAPEXsales}_{it} + \beta_7 \text{LTA2}_{it} + \beta_8 \lambda_{it} + \beta_9 \text{Industry\_dummies}_{it}$$

$$+ \beta_{10} \text{Year\_dummies}_{it} + v_{it} \tag{10}$$

where i identifies each firm, t indicates the year of observation (from 1 to 13), α and β_p are the coefficients to be estimated, and v_{it} represents the random disturbance for each observation. The dependent variable is excess value (*ExcessValue*), calculated following the Berger and Ofek (1995) imputed value approach, based on comparing the firm's market value against the estimated value the firm would have if all its divisions operated as individual entities (imputed value).[13]

We then estimate Equation (11) to test our hypothesis regarding the effect of the diversification pattern on a firm's value. The explanatory variable is our proposed index *DIVPAT(THEO)*, which measures RO patterns of diversification:

---

[12]An alternative approach to the Heckman two-step estimator is the Heckman maximum likelihood (ML) estimator. Whereas in the former, the selection equation and the outcome equation are estimated separately by probit and OLS estimations, respectively, in the Heckman ML estimator, both equations are estimated jointly in a single step by maximum likelihood. Assumptions for applying this ML approach are more restrictive than those required by the Heckman two-step estimator.

[13] If the excess value is negative, a discount will emerge, with diversification proving to be a value-destroying strategy. In contrast, a positive excess value will imply that the diversifier trades at a premium over its single-segment counterparts, a diversification strategy thus contributing towards enhancing a firm's value. See Berger and Ofek (1995) for more details. We calculate the "excess value" by dividing the enterprise's value by its imputed value, and then taking the natural logarithm of this ratio. Following Campa and Kedia's (2002) study, we compute a firm's market value (MV) as the sum of market value of equity (MVE), long-term (LtD), short-term (StD) debt, and preferred stock (PrefStock).

$$\text{EXCESS\_VALUE}_{it} = \alpha + \beta_1 \text{ DIVPAT(THEO)}_{it} + \beta_2 \text{ LTA}_{it} + \beta_3 \text{ LDTA}_{it} + \beta_4 \text{ EBITsales}_{it}$$

$$+ \beta_5 \text{ CAPEXsales}_{it} + \beta_6 \text{ LTA2}_{it} + \beta_7 \lambda_{it} + \beta_8 \text{ Industry\_dummies}_{it}$$

$$+ \beta_9 \text{Year\_dummies}_{it} + v_{it} \tag{11}$$

where i identifies each firm, t indicates the year of observation (from 1 to 13), α and β_p are the coefficients to be estimated, and v_{it} represents the random disturbance for each observation.

In line with prior research (Berger and Ofek, 1995; Campa and Kedia, 2002), we control for several firm-characteristics likely to impact excess value: firm size (*LTA*) and its squared term (*LTA2*), financial leverage (proxied by the ratio of long-term debt to total assets, *LDTA*), profitability (*EBITsales*), and level of investment (*CAPEXsales*). Following Santaló and Becerra (2008), we also incorporate the industry effect (*Industry_dummies*)[14]. Additionally, we control for the year effect (*Year_dummies*) and self-selection (λ). The estimated coefficient associated with the λ term is a key point in the analysis. A significant λ coefficient will mean that the correlation between the residuals of the selection equation and the outcome equation cannot be assumed to be zero, confirming the existence of selectivity.

Table 3 provides a summary of the descriptive statistics of the variables involved in the analysis for the sample[15]. Particularly noteworthy is the negative sign for the average excess value (-0.0574) reflecting a diversification discount. As regards the RO diversification pattern, companies display a balanced average position, which is not strongly inclined towards either of the two extreme strategies.

TABLE 3 ABOUT HERE

---

[14] Industry dummies are calculated at the 2-digit SIC code level.

[15] Summary statistics disaggregated by diversification status are available upon request from the contact author.

4. Empirical findings

4.1. Propensity to diversify: a probit estimation of the selection equation

Table 4 contains the probit estimation for the selection equation (Equation 7) as the first step in the Heckman method. Estimations in Columns (2) to (4) extend probit specification (1) by incorporating lags and year dummies. Goodness-of-fit (Pseudo-R^2) ranges between 0.15 and 0.16, comparable to prior literature. Among firm-characteristics, the level of investments (*CAPEXsales*) shows a negative and significant coefficient in all estimations. This result suggests that companies with low investment levels are more prone to diversify. Firm size (*LTA*) and its lag have a positive and highly significant coefficient, indicating that larger companies are more likely to incorporate multiple business units. Finally, firm profitability (*EBITsales*) is only statistically significant in the models where lagged variables are omitted. Our results evidence that less profitable enterprises are more liable to engage in this strategy.

TABLE 4 ABOUT HERE

Results concerning the effect of industry factors on the propensity to diversify are robust to the alternative estimations. Consistent with Campa and Kedia (2002) and Villalonga (2004b), results yield evidence that a greater presence of diversified firms in the core industry positively impacts the decision to diversify. As far as macroeconomic variables are concerned, they have no significant impact on diversification likelihood, *CONTRACTION* being only borderline significant in Column (3). There is also weak evidence concerning the relevance of the *changeGDP* variable. In the probit specification in Column (1), *changeGDP* is positively associated with the diversification decision, suggesting that companies are more

18

likely to diversify during cycles of economic growth. However, this variable does not retain its statistical significance in the remaining specifications.

In sum, we find that characteristics at firm-level and at industry-level are the main drivers of the diversification decision. Moving on to the second step of Heckman's approach and performing the estimations of our outcome equations, we take the specification of the selection equation in Column (1). In this way, we exclude lagged values of firm variables and time dummies which lack statistical significance in most cases, while minimising loss of observations for subsequent analyses. This probit ensures at least four exclusion restrictions since the variables *PNDIV*, *PSDIV*, *changeGDP* and *CONTRACTION* are included in the selection equation but not in the outcome equations, thus mitigating potential collinearity problems.

4.2. Components of the diversification pattern index and firm value

As preliminary analyses, we test the impact each dimension of our index has on a firm's value, both separately and jointly. Results are displayed in Table 5. The estimated parameters of Equation (8), which analyses the value effect of the inequality in the firm's level of participation in its businesses (*INTER*), are reported in Columns (1) and (2). We find clear evidence of a positive effect (statistically significant at the 1% level) of *INTER* on excess value. This result supports our hypothesis, since the greater the inequality in the distribution of business activity, the closer the firm's strategy to an OD. Columns (3) and (4) show the relationship between excess value and the company's overall commitment in all its business segments (*INTRA*) as defined in Equation (9).We find that *INTRA* is negatively associated (p-value=0.000) with excess value. This result is consistent with our arguments since holding major commitments in many businesses is negatively related to an OD. Finally, Columns (5) and (6) present the coefficient estimates for Equation (10) which accounts for both

19

dimensions (*INTER* and *INTRA*) simultaneously. Results are robust, with both dimensions of our index maintaining statistical significance above the 1% level.

TABLE 5 ABOUT HERE

Results are robust after dropping the extreme excess value observations (above 1.386 or below -1.386) from the sample (Berger and Ofek, 1995) as shown in Columns (7) to (12) of Table 5, and after including industry and year dummies (even columns in Table 5). Moreover, as can be seen, the λ coefficient is strongly statistically significant in all regressions (except for the estimations of Equation (8)), thus allowing us to reject the null hypothesis that the correlation between the residuals of the selection equation and the outcome equation is zero. This evidence confirms that our sample suffers from self-selection bias and thus Heckman's approach is justified. Furthermore, as indicated by the Wald test reported at the bottom of the tables, variables are jointly significant above the 1% level in all models.

4.3. Diversification pattern and firm value

Columns (1) and (2) in Table 6 provide interesting insights into the relevance of the diversification pattern for explaining value outcome (Equation (11)). Our research hypothesis receives strong support, since our evidence shows that *DIVPAT(THEO)* explains part of the diversification discounts/premiums. We find a significant (at the 1% level) and positive effect of *DIVPAT(THEO)* on excess value, indicating that a strategy further away from the 'extreme' AiPD profile implies higher excess values. Likewise, as a firm's diversification approaches an OD strategy (as measured by a longer Euclidean distance, thus higher *DIVPAT(THEO)*), diversification becomes a more value-enhancing strategy. These results suggest the value of a diversification strategy aimed not only at exploiting but also at seeding

new opportunities in further businesses. Results are robust to the exclusion of extreme excess value observations from the sample as shown in Columns (3) and (4) of Table 6, as well as to the addition of industry and year dummies (even columns of Table 6).

TABLE 6 ABOUT HERE

Finally, with regard to our control variables, only *LTA*, *EBITsales*, and *CAPEXsales* show any statistical significance above the 1% level in all estimations, all of them displaying a positive impact on excess value consistent with prior studies. The Wald test indicates that variables display joint significance in all regressions. In the vast majority of regressions, the λ coefficient contains statistical significance, even above 1% in certain cases, thus confirming the existence of self-selection bias in the sample.

4.4. Robustness analyses

We conduct a number of robustness tests. First, we define the extreme AiPD profile of reference by using calibration samples based on the median scores of each dimension of our index for a subsample of companies with lower growth opportunities. Second, following prior literature, we examine the consistency of our empirical findings when controlling for a firm's diversification scope and growth opportunities. We then go further in testing our hypothesis and evaluate the robustness of our results when considering the risk of pre-emption. Finally, we control for the dynamic nature of the investment pattern by including several lags of our index.

4.4.1. Controlling for an alternative specification of the reference profile

Rather than predetermining the extreme AiPD profile theoretically as detailed earlier, we re-estimate our Equation (11) results by defining our diversification pattern index based on an

21

AiPD profile of reference empirically specified. To achieve this aim, we use a calibration sample (Venkatraman, 1989; Venkatraman and Prescott, 1990), comprising the bottom ten per cent of firms according to growth opportunities[16]. Growth opportunities are proxied by either Tobin's Q (Cao, Simin and Zhao, 2008) or the ratio of R&D expenses to total sales (Mehran, 1995). The AiPD reference point is determined by the median scores along both index dimensions (*INTER* and *INTRA*) for the calibration sample. Table 7 reports a statistical summary of both dimensions for this subsample.

TABLE 7 ABOUT HERE

We construct alternative proxies for the diversification profile index depending on the growth opportunities proxy used for calibrating. The vector of scores for the AiPD extreme profile is then replaced in Equation (5) by the median scores of *INTER* and *INTRA* in the calibration sample, this sample being determined by either Tobin's Q (*Q*) or the R&D expenses to total sales (*RDsales)* ratio. The redefined diversification pattern indexes are denoted by *DIVPAT(CAL_Q)* and *DIVPAT(CAL_RD),* respectively[17].

Results are displayed in Table 8. We find that our results are robust to the alternative specifications of the reference pattern (*DIVPAT(CAL_Q)* in Column (1), and *DIVPAT(CAL_RD)* in Column (3) in Table 8), the elimination of extreme excess value observations (Columns (5) to (8) of Table 8), the inclusion of industry and year dummies (even columns in Table (8)), as well as Heckman's ML estimation[18].

---

[16]The most extreme cases of AiPD should imply the lowest growth opportunities values as this pattern is primarily aimed at exercising options in one full-scale step.

[17] Following Vekatraman and Prescott (1990), in regressions where indexes based on the calibration samples are used, both the bottom (calibration sample) as well as the top ten percent of firms according to their level of growth opportunities (proxied either by *Q* or *RDsales*) are excluded from the study sample.

[18] Results of this paper are also robust to the alternative Heckman ML approach. All Heckman's ML estimations are omitted in the interest of saving space. Results are available upon request from the corresponding author.

TABLE 8 ABOUT HERE

4.4.2. Controlling for the diversification status

We further check the robustness of the results for Equation (11) by restricting the study sample to diversified firms and then estimating by OLS[19]. Table 9 displays our results. Odd-numbered Columns (1), (3) and (5) in Table 9 report the results from the estimation of this equation with the alternative specifications of the diversification pattern index (*DIVPAT(THEO)*, *DIVPAT(CAL_Q)* and *DIVPAT(CAL_RD)*, respectively); and the even-numbered Columns (2), (4) and (6) show re-estimated coefficients when including time dummies and industry dummies. Overall, our findings also hold when Equation (11) is estimated only on the subsample of diversified firms. These results are also robust to the exclusion of extreme excess value observations from the sample (Columns (7) to (12) in Table 9).

TABLE 9 ABOUT HERE

Alternatively, we redefine Equation (11) to include the level of diversification as a control variable, approximated by the modified Herfindahl index ($MHERF = 1 - \sum P_h^2$):

$$EXCESS\_VALUE_{it} = \alpha + \beta_1 \, DIVPAT(THEO)_{it} + \beta_2 \, MHERF_{it} + \beta_3 \, LTA_{it} + \beta_4 \, LDTA_{it}$$

$$+ \beta_5 \, EBITsales_{it} + \beta_6 \, CAPEXsales_{it} + \beta_7 \, LTA2_{it}$$

$$+ \beta_8 \, Industry\_dummies_{it} + \beta_9 \, Year\_dummies_{it} + v_{it} \qquad (12)$$

Table 10 reports these additional sensitivity tests. Column (1) presents estimation results for the core model of Equation (12) only with control variables. Columns (2) to (4) show

---

[19] As unisegment firms are excluded, there is no reason to control for selectivity.

results for alternative specifications of the diversification pattern index (*DIVPAT(THEO)*, *DIVPAT(CAL_Q)* and *DIVPAT(CAL_RD)*, respectively) when simultaneously adding the scope dimension (*MHERF*). Columns (5) to (8) evidence the consistency of these results after dropping extreme excess values from the sample. In line with prior literature, our sample also shows a diversification discount, as displayed in Columns (1) and (5) of Table 10. Results concerning the pattern index are robust across all estimations, once again bearing out that the closer a firm's diversification profile is to an OD, the higher the excess value. Interestingly, once the pattern of diversification is accounted for in the regressions, the documented discount becomes a premium, which is statistically significant in regressions where extreme excess value observations are not excluded (Columns (2) to (4)). It appears that the conflicting evidence regarding the impact of diversification on firm value may partly be explained by the fact that prior analyses might be mixing the effects of different dimensions of diversification, namely scope and investment pattern. This may require proper separate identification and measurement to investigate the overall impact of diversification more accurately.

TABLE 10 ABOUT HERE

4.4.3. *Controlling for the diversification status and growth opportunities*

A further robustness check is to run our regressions controlling for a firm's growth opportunities together with diversification level (Equation (13)). Following Grullon, Lyandres and Zhdanov (2012), growth opportunities (*GROWTH*) are alternatively proxied by: firm age (*lnAge*, defined as natural logarithm of the difference between the current year and the founding year of the firm) and R&D intensity (defined as the ratio of annual R&D expenditures and beginning-of-year book assets, *RDA*.

24

$$\text{EXCESS\_VALUE}_{it} = \alpha + \beta_1 \text{ DIVPAT(THEO)}_{it} + \beta_2 \text{ MHERF}_{it} + \beta_3 \text{ lnAge}_{it} + \beta_4 \text{ LTA}_{it}$$

$$+ \beta_5 \text{ LDTA}_{it} + \beta_6 \text{ EBITsales}_{it} + \beta_7 \text{ CAPEXsales}_{it} + \beta_8 \text{ LTA2}_{it}$$

$$+ \beta_9 \text{ Industry\_dummies}_{it} + \beta_{10} \text{ Year\_dummies}_{it} + v_{it} \qquad (13)$$

Table 11 summarises the results of Equation (13) with the alternative specifications of our diversification pattern index. Columns (1) to (3) use *lnAge* as growth opportunities proxy, while Columns (4) to (6) are based on *RDA*. Only *RDA* displays statistical significance. Overall, our hypothesis also receives strong support after controlling for growth opportunities in the regressions. Across the alternative index specifications, the diversification profile proves to be statistically significant (above 1% level) to explain diversification discounts/premiums. Its coefficient displays a positive sign which again suggests that a further deviation (in terms of Euclidean distance) from an AiPD strategy, and thereby a closer profile to an OD strategy, is more value-enhancing.

TABLE 11 ABOUT HERE

4.4.4. *Controlling for pre-emption risk*

To explore the possible effect of industry heterogeneity on a firm's pattern of diversification, we control for the threat of pre-emption. In intensively competitive industries, such sequential small-scale investments which characterise OD may prove counter-valuable. They may imply the loss of first-mover advantage and give competitors a clue about the strategy followed-up by the firm before a major commitment to secure such a potential competitive position has been made, thereby reducing the growth option lifespan and its value. As a result, we expect an OD to be less valuable the higher the risk of pre-emption. Following Folta and Miller (2002: 83), the risk of preemption is approximated by the number

of rivals actively operating in the same product domain. We gather yearly data on the total number of U.S. firms by NAICS codes from the U.S. Census Bureau and then match NAICS codes with SIC codes. Our variable to proxy risk of pre-emption is *PREEMPT*, calculated as the natural logarithm of the number of firms operating in the same 2-digit SIC code industry as the core business of the corresponding firm.

We split our full sample into terciles of *PREEMPT* levels and re-estimate Equation (12). Table 12 displays the coefficients of the Heckman second-stage estimations taking the alternative specifications of diversification pattern index. Across all subsamples, the coefficient on the diversification pattern is positive and statistically significant at the 1% level, thus confirming our hypothesis that an OD strategy leads to greater excess values. Interestingly, this coefficient is not only statistically significant but also smaller the higher the risk of pre-emption in the industry. For example, the coefficient on *DIVPAT(THEO)* decreases almost 44% within the subsample of firm-years in the third tercile of *PREEMPT* compared to the first tercile. Consistent with our arguments, a higher threat of preemption by competitors detracts value from an OD strategy since a competitor may more easily snatch a business opportunity open to the firm but to which it is not yet fully committed. Our results concur with prior RO literature showing the detrimental effects of the risk of pre-emption of growth option value in a variety of contexts, such as licensing (Jiang, Aulakh and Pan, 2009), or buyouts in equity partnerships (Folta and Miller, 2002)).

TABLE 12 ABOUT HERE

4.4.5. *Controlling for the dynamic effect*

A firm's diversification pattern may shape and evolve over time as a result of different investment decisions at multiple time points. This suggests the possible influence of past

values of the index on current diversification outcomes. In a dynamic context, diversification investment strategies, especially those closer to an RO logic, are expected to be highly variable across time, as growth options are continuously explored and exploited[20]. Thus, we extend our core empirical model to a dynamic setting by incorporating the effect of lagged *DIVPAT(THEO)*:

$$
\begin{aligned}
\text{EXCESS\_VALUE}_{it} = {} & \alpha + \beta_1 \, \text{DIVPAT(THEO)}_{it} + \beta_2 \, \text{DIVPAT(-1)}_{it} + \beta_3 \, \text{DIVPAT (-2)}_{it} \\
& + \beta_4 \, \text{MHERF}_{it} + \beta_5 \text{LTA}_{it} + \beta_6 \, \text{LDTA}_{it} + \beta_7 \, \text{EBITsales}_{it} \\
& + \beta_8 \, \text{CAPEXsales}_{it} + \beta_9 \, \text{LTA2}_{it} + \beta_{10} \, \text{Industry\_dummies}_{it} \\
& + \beta_{11} \, \text{Year\_dummies}_{it} + \nu_{it}
\end{aligned}
\tag{14}
$$

Table 13 reports the Heckman two-step estimation results of Equation (14). Column (1) shows the results with one lag of *DIVPAT(THEO)* (denoted by *DIVPAT(-1)*), and Column (2) estimates the full model with the two-lagged *DIVPAT(THEO)* variable (represented as *DIVPAT(-2)*). Our primary results remain similar. Evidence presented in both columns confirms that the diversification pattern is statistically significant to explain the excess value. Consistent with our hypothesis, we find that excess value is higher the greater the *DIVPAT(THEO)*, in other words, the closer the firm is to an OD strategy. Results also support the dynamic nature of this relationship. The one-lagged index *DIVPAT(-1)* is positive and statistically significant at the 1% level, meaning that any increase in this variable (and thereby, greater proximity to an OD strategy) will produce an increase in subsequent excess values. Interestingly, we find that the economic significance decreases over time, thus

---

[20] Following one referee's suggestion, we estimated the volatility of each component of our index (*INTER* and *INTRA*), in terms of standard deviation over the sample years within each firm. We split the sample into two groups according to the values of our diversification pattern index *DIVPAT(THEO)* to identify those firms with a diversification strategy closer to an AiPD (below-median subsample) and those companies with a diversification strategy closer to an OD (above-median subsample). We find that volatility is higher in the above-median subsample, both in mean and median terms, confirming that OD strategies tend to be more dynamic across time. Results are available upon request from the contact author.

suggesting that the effect of past diversification profiles attenuates over time. The current value of diversification pattern drives the main effect since it not only builds on the diversification history of the company but especially on the current configuration of the firm's growth options across its portfolio of businesses.

TABLE 13 ABOUT HERE
----

5. Discussion and conclusion

This paper investigates whether the pattern of diversification, which entails a different configuration of the firm's growth options portfolio, accounts for part of the diversification discounts/premiums. We approximate each firm's strategy position relative to two contrasting diversification paths (ranging from an AiPD to an OD one) and directly analyse the effect of the investment pattern on diversification premium/discount measures. We perform our analysis on a sample of U.S. firms from 1998 to 2010, offering updated evidence on a post-1997 sample, after implementation of the new SFAS 131 reporting standard in the U.S. Our results confirm that how the firm diversifies is by no means a trivial issue when determining diversification value outcomes.

This evidence supports the basic premise by Williamson (2001) concerning the relevance of creating strategic options for the future. Results reveal that firms are likely to enhance their value if they spread their current capabilities beyond their core businesses by engaging in underdeveloped participations in new industries for strategic or explorative purposes. This investment logic is in line with RO rationale, and emphasises the importance of having *"a foot in the door"* to access future investment opportunities, enriching the firm's set of growth

28

options, whilst at the same time limiting downside risk by delaying full commitment of resources in an effort to capitalise on uncertainty.

Some important contributions for diversification literature emerge from our empirical findings. Firstly, our results suggest considering an additional dimension apart from the commonly studied dimensions of diversification (scope and relatedness): the diversification pattern seems to moderate the diversification-value linkage in a way which is consistent with the real options logic. Our findings tie in with recent streams of research which advocate the endogenous nature of the diversification decision, thus making value creation or destruction contingent on firm-specific characteristics rather than on generic characteristics attached to the strategy or the firms undertaking it. We provide evidence that diversification strategy is neither intrinsically good nor bad. Rather, our findings suggest that when exploring the *diversification puzzle*, what seems important is not only how much to diversify (scope) and where (relatedness between businesses), but also "how". Interestingly, our study sheds light on the need to explore further dimensions of diversification. In this paper, we study the diversification pattern and find that it accounts for the diversification discount/premium. Failing to consider different sides in this strategy may have given rise to such conflicting evidence in prior literature as a result of mixing the different dimensions of diversification, each of which has a different impact on a firm's value.

Furthermore, many papers call for the need to investigate further the validity of real options for strategic analysis in an effort to advance theory (Reuer and Tong, 2007). This study contributes to filling the gap in empirical works which apply the RO approach to strategy. We adopt this fresh approach to join a long-standing debate on the diversification-value relationship, which still awaits a suitable answer, as evidenced by recent works (Elsas *et al.*, 2010; Hoechle *et al.*, 2012; Andreou *et al.*, 2016; Kuppuswamy and Villalonga, 2016).

We have shown that such a debate could benefit from linking RO based patterns of diversification and firm value.

Finally, we develop and test a proxy for the diversification profiles drawn on the way growth opportunities and assets-in-place are handled. To the best of our knowledge, this constitutes the first attempt to capture and measure this dimension of corporate diversification. Our two-dimensional index follows the notion of *profile deviation* (Venkatraman, 1989), which has been applied in other areas in strategy, and joins prior research works analyzing the development of additional indexes to better characterize each firm's diversification strategy, such as Jacquemin and Berry's (1979) entropy index or Neffke and Henning's (2013) skill-relatedness index.

This study has significant implications for business management since we show managers how important the way in which such a strategy is implemented may prove to be vis-à-vis value creation. We advocate proactive managerial behaviour and stress the vital importance of combining expansion in a firm's core segments with the simultaneous opening of fresh strategic options in new businesses. Our results reveal that the pattern of diversification has a significant impact on a firm's market value.

Finally, we point to certain limitations in our research and to questions which remain for future study. First, our sample only comprises U.S. firms. It might prove interesting to replicate the analysis on an international sample and check the consistency of our results. Secondly, our empirical analysis may inevitably suffer from the limitation of using segment data, which can sometimes be biased due to inconsistent segment reporting decisions or managerial self-interested reporting changes that do not correspond to an actual change in business operations (Villalonga, 2004a). In addition, further research should focus on dealing with the diversification-value relationship in the current financial crisis. The implications of

both AiPD and OD for corporate value may become more marked in a context of crisis, thus reinforcing the moderating role of the diversification pattern in the diversification-value relationship. Finally, our analysis suggests it may prove interesting to redefine and amplify our index to a dynamic context so as to control further for the impact of its evolution on value.

Acknowledgements

References

Amihud, Y., Lev, B., 1981. Risk reduction as a managerial motive for conglomerate mergers.

31

Bell Journal of Economics 12(2), 605–617.

Andreou, P., Louca, C., Petrou, A., 2016. Organizational learning and corporate diversification performance. Journal of Business Research, *forthcoming*.

Berger, P. G., Ofek, E., 1995. Diversification's effect on firm value. Journal of Financial Economics 37(1), 39-65.

Bernardo, A., Chowdhry, B., 2002. Resources, real options, and corporate strategy. Journal of Financial Economics 63(1), 211-234.

Borghesi, R., Houston, J., Naranjo, A., 2007. Value, survival and the evolution of firm organizational structure. Financial Management 36(3), 5-31.

Bowman, E. H., Hurry, D., Miller, A., 1992. Calls on high technology: Japanese exploration of venture capital investments in the United States. Strategic Management Journal 13(2), 85-101.

Campa, J. M., Kedia, S., 2002. Explaining the diversification discount. The Journal of Finance 57(4), 1731-62.

Cao, C., Simin, T., Zhao, J., 2008. Can growth options explain the trend in idiosyncratic risk?. The Review of Financial Studies 21(6), 2599-2633.

Elsas, R., Hackethal, A., Holzhäuser, M., 2010. The anatomy of bank diversification. Journal of Banking & Finance 34(6), 1274-1287.

Fauver, L., Houston, J., Naranjo, A., 2003. Capital market development, international integration, legal systems, and the value of corporate diversification: A cross-country analysis. Journal of Financial and Quantitative Analysis 38(1), 135-157.

Ferris, S.P., Sen, N., Lim, C.Y., Yeo, G., 2002. Corporate focus versus diversification: the role of growth opportunities and cash flow. Journal of International Financial Markets

Institutions and Money 12(3), 231–252.

Folta, T. B., Miller, K.D., 2002. Real options in equity partnerships. Strategic Management Journal 23(1), 77-88.

Greene, W., 2003. Econometric analysis. Ed. Prentice Hall, United States, Fifth edition.

Grullon, G., Lyandres, E., Zhdanov, A., 2012. Real options, volatility, and stock returns. The Journal of Finance 67(4), 1499-1537.

Heckman, J., 1979. Sample selection bias as a specification error. Econometrica 47(1), 153-161.

Hirschman, A.O., 1964. The paternity of an index. American Economic Review 54(5), 761-770

Hoechle, D., Schmid, M., Walter, I., Yermack, D., 2012. How much of the diversification discount can be explained by poor corporate governance? Journal of Financial Economics 103(1), 41-60.

Holder, M., Zhao, A., 2015. Value exploration and materialization in diversification strategies. Review of Quantitative Finance and Accounting 45(1), 175-213.

Jacquemin, A., Berry, C., 1979. Entropy measure of diversification and corporate growth. The Journal of Industrial Economics 27(4), 359-369.

Jiang, M. S., Aulakh, P. S., Pan, Y., 2009. Licensing duration in foreign markets: a real options perspective. Journal of International Business Studies 40(4), 559-577.

Klingebiel, R., Adner, R, 2015. Real options logic revisited: The performance effects of alternative resource allocation regimes. Academy of Management Journal 58(1), 221–241.

Kogut, B., Kulatilaka, N., 1994. Options thinking and platform investments: investing in opportunity. California Management Review 36(2), 52-71.

Kuppuswamy, V., Villalonga, B., 2016. Does diversification create value in the presence of external financing constraints? Evidence from the 2007-2009 financial crisis. Management Science, *forthcoming*.

Lang, L., Stulz, R. M., 1994. Tobin's q, corporate diversification and firm performance. Journal of Political Economy 102(6), 1248-80.

Martin, J.D., Sayrak, A., 2003. Corporate diversification and shareholder value: a survey of recent literature. Journal of Corporate Finance 9(1), 37–57.

Mehran, H., 1995. Executive compensation structure, ownership, and firm performance. Journal of Financial Economics 38(2), 163–184.

Miller, D. J., 2006. Technological diversity, related diversification, and firm performance. Strategic Management Journal 27(7), 601-619.

Mitton, T., Vorkink, K., 2010. Why do firms with diversification discounts have higher expected returns? Journal of Financial and Quantitative Analysis 45(6), 1367-1390.

Neffke, F., Henning, M., 2013. Skill relatedness and firm diversification. Strategic Management Journal 34(3), 297-316.

Palich, L.E., Cardinal, L.B., Miller, C., 2000. Curvilinearity in the diversification-performance linkage: an examination of over three decades of research. Strategic Management Journal 21(2), 155-174.

Rajan, R., Servaes, H., Zingales, L., 2000. The cost of diversity: the diversification discount and inefficient investment. The Journal of Finance 55(1), 35-80.

Raynor, M. E., 2002. Diversification as real options and the implications of firm-specific risk and performance. The Engineering Economist 47(4), 371-389.

Reuer, J.J., Tong, T.W., 2007. How do real options matter? Empirical research on strategic

investments and firm performance, in: Reuer, J.J., Tong, T.W. (Eds), Real Options Theory. Advances in Strategic Management. Volume 24, Elsevier, First edition.

Rudolph C., Schwetzler, B., 2013. Conglomerates on the rise again? A cross-regional study on the impact of the 2008–2009 financial crisis on the diversification discount. Journal of Corporate Finance 22, 153–165.

Santaló, J., Becerra, M., 2008. Competition from specialized firms and the diversification-performance linkage. The Journal of Finance 63(2), 851-883.

Servaes, H., 1996. The value of diversification during the conglomerate merger wave. The Journal of Finance 51(4), 1201-1225.

Stowe, J. D., Xing, X., 2006. Can growth opportunities explain the diversification discount?. Journal of Corporate Finance 12(4), 783-796.

Teplensky, J. D., Kimberly, J. R., Hillman, A. L., Schwartz, J. S., 1993. Scope, timing and strategic adjustment in emerging markets: manufacturer strategies and the case of MRI. Strategic Management Journal 14(7), 505-527.

Trigeorgis, L., Lambertides, N., 2014. The role of growth options in explaining stock returns. Journal of Financial and Quantitative Analysis 49(3), 749-771.

Venkatraman, N., 1989. The concept of fit in strategy research: toward verbal and statistical correspondence. The Academy of Management Review 14(3), 423-444.

Venkatraman, N., Prescott, J.E., 1990. Environment-strategy coalignment: an empirical test of its performance implications. Strategic Management Journal 11(1), 1-23.

Villalonga, B., 2004a. Diversification discount or premium? New evidence from BITS establishment-level data. The Journal of Finance 59(2), 479-506.

Villalonga, B., 2004b. Does diversification cause the "diversification discount"?. Financial Management 33(2), 5-27.

Williamson, P., 2001. Strategy as options on the future, in: Cusumano, M. and Markides, C. (eds.), Strategic Thinking for the Next Economy, MIT Sloan Management Review, Ed. Jossey-Bass A Wiley Company, Cambridge, MA, Chapter 7, pp. 157-178

Table 1

Distribution of observations in the sample (SIC codes classification)

YEAR	INITIAL SAMPLE (Obs.)					FINAL SAMPLE (Obs.)		
	UNI	MULTI	TOTAL	Financial industry	Does not meet other criteria	UNI	MULTI	TOTAL
1998	4282	1726	6008	1603	2392	1243	770	2013
1999	4309	1842	6151	1607	2433	1247	864	2111
2000	4083	1902	5985	1548	2326	1206	905	2111
2001	3977	1841	5818	1478	2311	1133	896	2029
2002	4133	1643	5776	1517	2250	1205	804	2009
2003	4337	1495	5832	1550	2283	1271	728	1999
2004	4506	1394	5900	1490	2308	1403	699	2102
2005	4670	1318	5988	1477	2338	1500	673	2173
2006	4746	1251	5997	1450	2238	1655	654	2309
2007	4719	1159	5878	1404	2082	1767	625	2392
2008	4479	1096	5575	1345	1807	1825	598	2423
2009	4319	1060	5379	1304	1853	1688	534	2222
2010	4067	983	5050	1214	1523	1787	526	2313
TOTAL	56627	18710	75337	18987	28144	18930	9276	28206
%	75.165%	24.835%	100.000%			67.113%	32.887%	100.000%

This Table shows the distribution of the firm-year observations between unisegment and diversified firm subsamples, before and after applying Berger and Ofek's (1995) sample selection criteria. The central columns contain the number of observations dropped from the initial sample due to the report of any business segment in the financial sector or failure to meet the other Berger and Ofek criteria.

Table 2

Descriptive statistics of the data [U.S. companies (1998-2010)]

Variable	N	Mean	Median	STD	Min.	Max.	1st quartile	3rd quartile
Total Sales	16554	1475.991	308.7555	4265.445	20.01	98540	90.933	1120.056
Total Assets	16554	1400.56	320.502	2920.466	4.5800	21972	91.9698	1187.725
Common Equity	16554	612.1729	172.485	1250.963	0.2387	15835	52.312	556.082
EBIT	16554	113.9356	19.1145	352.1681	-6740.195	5039	1.522	94.724
Market capitalization	16554	1731.068	353.8434	4482.616	1.3400	78973.82	88.0205	1282.935
Total Debt	16554	367.2181	31.167	959.7748	0	12358.83	2.309	261.523
Total observations	16,554 observations after removing outliers							
	[Unisegment firms: 12,047 obs. (72.77%); diversified firms: 4,507 obs. (27.23%)]							

This Table contains descriptive statistics of general financial variables for the final sample of 16,554 firm-year observations, for both unisegment (12,047 firm-year observations) and multisegment companies (4,507 firm-year observations). Figures are expressed in million US$.

Table 3
Summary statistics of variables for the full sample (1998-2010)

Variable	N	Mean	Median	STD	Min.	Max.	1st quartile	3rd quartile
Excess Value	16554	-0.0574	0.0000	0.7875	-2.8458	2.6628	-0.5338	0.4335
Excess Value (without extremes EV)	15104	-0.0141	0.0000	0.6126	-1.3846	1.3858	-0.4414	0.4113
INTER	16554	0.1554	0	0.2986	0	0.9999	0	0.1039
INTRA	16554	0.5037	0.5000	0.4692	0	1	0	1
DIVPAT(THEO)	16554	0.5644	0.8710	0.5111	0	1.4141	0	1
Control variables								
LTA	16554	5.8406	5.7699	1.7308	1.5217	9.9975	4.5215	7.0798
EBITsales	16554	0.0543	0.0681	0.1843	-1.1784	1.1792	0.0143	0.1303
CAPEXsales	16554	0.0684	0.0332	0.1105	0	0.9348	0.0166	0.0677
LDTA	16554	0.1581	0.1187	0.1617	0	0.7391	0.0016	0.2687
PNDIV	16554	0.4364	0.4231	0.2194	0	1	0.2857	0.5714
PSDIV	16554	0.5549	0.5919	0.2973	0	1	0.3325	0.7960
changeGDP	16554	0.0222	0.0270	0.0195	-0.0260	0.0480	0.0180	0.0360
CONTRACTION	16554	1.6651	0	3.0931	0	9	0	0

This Table shows descriptive statistics of the variables involved in our models for the final sample of 16,554 firm-year observations of unisegment (12,047 firm-year observations) and multisegment companies (4,507 firm-year observations). Some observations contain missing data for certain variables.
Excess Value is the measure developed by Berger and Ofek (1995) to assess the value created by diversifying. *INTER* and *INTRA* are the two-dimensional index component variables. *DIVPAT(THEO)* is the diversification pattern index based on the Euclidean distance and with the extreme AiPD pattern of reference theoretically specified. Control variables: *LTA* (size), *EBITsales* (profitability), *CAPEXsales* (level of investment in current operations), *LDTA* (financial leverage), *PNDIV* (fraction of firms in the firm's core industry that are diversified), *PSDIV* (the proportion of the firm's core industry sales accounted for by diversifiers), *changeGDP* (real growth rates of gross domestic product), *CONTRACTION* (the number of months in the year the U.S. economy was in recession). Figures are expressed in million US$.

Table 4

Firms' propensity to diversify [first stage of Heckman's estimation] (Eq. 7)

PROBIT	(1) baseline model	(2) with lags	(3) with dummy years	(4) with lags and dummy years
Constant	-2.6356***	-2.6479***	-2.7298***	-2.7800***
	(0.0602)	(0.0732)	(0.0713)	(0.0850)
Firm characteristics				
LTA	0.1108***	-0.0474	0.1134***	-0.0470
	(0.0069)	(0.0658)	(0.0069)	(0.0661)
EBITsales	-0.1592**	-0.0450	-0.1391**	-0.0192
	(0.0655)	(0.1075)	(0.0660)	(0.1083)
CAPEXsales	-0.7678***	-0.5510**	-0.7887***	-0.5599**
	(0.1115)	(0.2610)	(0.1122)	(0.2620)
LTA $_{t-1}$		0.1678***		0.1700***
		(0.0655)		(0.0658)
EBITsales $_{t-1}$		-0.0531		-0.0713
		(0.1018)		(0.1027)
CAPEXsales $_{t-1}$		-0.3274		-0.3357
		(0.2465)		(0.2482)
Industry characteristics				
PNDIV	2.1820***	2.1367***	2.1500***	2.1232***
	(0.0682)	(0.0810)	(0.0702)	(0.0835)
PSDIV	0.5856***	0.6358***	0.5770***	0.6252***
	(0.0492)	(0.0594)	(0.0495)	(0.0597)
Macroeconomic characteristics				
ChangeGDP	2.4047**	0.9054	1.4355	0.1500
	(0.9429)	(1.1328)	(1.4136)	(1.5901)
CONTRACTION	0.0074	-0.0049	0.0157*	0.0103
	(0.0058)	(0.0068)	(0.0084)	(0.0095)
Year dummies	NO	NO	YES	YES
N. of obs.	16554	11745	16554	11745
Log. Likelihood	-8177.981	-5755.1745	-8167.4142	-5746.8648
Pseudo-R^2	0.1562	0.1590	0.1573	0.1602

This Table shows probit estimation results for the selection equation (Equation 7) as the first stage of Heckman's procedure. The dependent variable takes the value 1 when the firm is diversified and zero otherwise. The pseudo-R square indicates the goodness of fit. Standard error is shown in parentheses under coefficients. ****, ** and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 5
The dimensions of the index and Excess Value [Heckman second stage - Eq. (8) to (10)]

	Model ExcessValue = f(INTER, INTRA, control variables)											
	Dependent variable: Excess Value						Dependent variable: Excess Value (without extremes)					
	INTER	INTER with industry and time dummies	INTRA	INTRA with industry and time dummies	INTER and INTRA	INTER and INTRA with industry and time dummies	INTER	INTER with industry and time dummies	INTRA	INTRA with industry and time dummies	INTER and INTRA	INTER and INTRA with industry and time dummies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
constant	-1.7733*** (0.1314)	-1.8529*** (0.1446)	-1.1200*** (0.1315)	-1.1677*** (0.1452)	-1.2203*** (0.1319)	-1.2820*** (0.1450)	-0.9891*** (0.1160)	-1.0721*** (0.1266)	-0.5802*** (0.1168)	-0.6532*** (0.1275)	-0.6713*** (0.1170)	-0.7508*** (0.1275)
INTER	0.1237*** (0.0375)	0.1630*** (0.0373)			0.2395*** (0.0371)	0.2871*** (0.0371)	0.1300*** (0.0318)	0.1394*** (0.0318)			0.2119*** (0.0320)	0.2257*** (0.0321)
INTRA			-0.6570*** (0.0416)	-0.6220*** (0.0418)	-0.7077*** (0.0422)	-0.6896*** (0.0424)			-0.3967*** (0.0361)	-0.3774*** (0.0363)	-0.4479*** (0.0368)	-0.4369*** (0.0370)
control variables												
LTA	0.3983*** (0.0408)	0.3504*** (0.0403)	0.3426*** (0.0399)	0.2995*** (0.0395)	0.3440*** (0.0397)	0.3028*** (0.0393)	0.1758*** (0.0359)	0.1444*** (0.0355)	0.1477*** (0.0356)	0.1184*** (0.0351)	0.1508*** (0.0354)	0.1242*** (0.0349)
LDTA	-0.1674** (0.0797)	-0.1165 (0.0804)	-0.1461* (0.0777)	-0.1348* (0.0787)	-0.1536** (0.0774)	-0.1395* (0.0781)	-0.1335** (0.0671)	-0.1449** (0.0682)	-0.1194* (0.0662)	-0.1573** (0.0675)	-0.1277* (0.0659)	-0.1621** (0.0671)
EBITsales	0.5982*** (0.0748)	0.6902*** (0.0733)	0.6856*** (0.0732)	0.7539*** (0.0719)	0.6593*** (0.0731)	0.7190*** (0.0716)	0.5231*** (0.0644)	0.5889*** (0.0637)	0.5900*** (0.0638)	0.6382*** (0.0631)	0.5697*** (0.0635)	0.6153*** (0.0629)
CAPEXsales	1.0004*** (0.1269)	1.7592*** (0.1473)	0.8943*** (0.1242)	1.6257*** (0.1447)	0.8197*** (0.1243)	1.5391*** (0.1443)	0.5956*** (0.1050)	1.1035*** (0.1240)	0.5632*** (0.1037)	1.0559*** (0.1228)	0.4947*** (0.1037)	0.9850*** (0.1225)
LTA2	-0.0220*** (0.0032)	-0.0175*** (0.0032)	-0.0202*** (0.0031)	-0.0160*** (0.0031)	-0.0203*** (0.0031)	-0.0164*** (0.0031)	-0.0071** (0.0028)	-0.0036 (0.0028)	-0.0064** (0.0028)	-0.0031 (0.0027)	-0.0067** (0.0028)	-0.0037 (0.0027)
Inverse Mills Ratio (λ_i)	-0.0197 (0.0286)	-0.0430 (0.0318)	-0.1664*** (0.0296)	-0.1545*** (0.0322)	-0.1842*** (0.0296)	-0.1725*** (0.0321)	0.0140 (0.0237)	-0.0068 (0.0267)	-0.0773*** (0.0249)	-0.0771*** (0.0273)	-0.0941*** (0.0249)	-0.0920*** (0.0273)
Industry dummies	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Year dummies	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
No. of Obs.	16554	16554	16554	16554	16554	16554	15104	15104	15104	15104	15104	15104
No. Censored Obs.	12047	12047	12047	12047	12047	12047	11007	11007	11007	11007	11007	11007
No. Uncensored Obs.	4507	4507	4507	4507	4507	4507	4097	4097	4097	4097	4097	4097
Wald Chi2 χ^2	692.47***	1243.87***	959.15***	1493.62***	1007.89***	1571.25***	456.62***	889.74***	569.80***	995.38	618.91***	1055.94***

This Table shows the Heckman second stage estimation by OLS (Heckman's two-step estimator) of Equations (8) to (10). The selection equation (Column 1 in Table 4) models firms' propensity to diversify. *ExcessValue* is regressed on *INTER*, *INTRA* and both dimensions together. Firm size (*LTA*) and its squared term (*LTA2*), financial leverage (*LDTA*), profitability (*EBITsales*), level of investment (*CAPEXsales*), industry effect (*industry dummies*), and time effect (*year dummies*) are controlled in the estimations. The Inverse Mills Ratio ((λ_i) is included as an additional regressor to correct potential self-selection bias in the sample. The Wald test contrasts the null hypothesis of no joint significance of the explanatory variables. Standard error is shown in parentheses under coefficients. ****, ** and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 6
Diversification pattern index (theoretically specified) and Excess Value [Heckman second stage - Eq. (11)]

	Model ExcessValue = f(DIVPAT(THEO), control variables)			
	Dependent variable: Excess Value		Dependent variable: Excess Value (without extremes)	
	Baseline index	Baseline index with industry and time dummies	Baseline index	Baseline index with industry and time dummies
	(1)	(2)	(3)	(4)
constant	-2.1036***	-2.1576***	-1.2241***	-1.3081***
	(0.1286)	(0.1410)	(0.1149)	(0.1252)
DIVPAT(THEO)	0.7592***	0.7936***	0.5354***	0.5497***
	(0.0470)	(0.0474)	(0.0413)	(0.0418)
control variables				
LTA	0.3583***	0.3126***	0.1560***	0.1283***
	(0.0398)	(0.0392)	(0.0353)	(0.0348)
LDTA	-0.1642**	-0.1380*	-0.1327**	-0.1620**
	(0.0776)	(0.0782)	(0.0659)	(0.0669)
EBITsales	0.6201***	0.6919***	0.5521***	0.6032***
	(0.0730)	(0.0714)	(0.0633)	(0.0626)
CAPEXsales	0.8097***	1.5463***	0.4854***	0.9822***
	(0.1245)	(0.1442)	(0.1036)	(0.1222)
LTA2	-0.0208***	-0.0168***	-0.0069**	-0.0039
	(0.0031)	(0.0031)	(0.0028)	(0.0027)
Inverse Mills Ratio (λ_i)	-0.1631***	-0.1602***	-0.0897***	-0.0902
	(0.0294)	(0.0319)	(0.0247)	(0.0271)
Industry dummies	NO	YES	NO	YES
Year dummies	NO	YES	NO	YES
No. of Obs.	16554	16554	15104	15104
No. Censored Obs.	12047	12047	11007	11007
No. Uncensored Obs.	4507	4507	4097	4097
Wald Chi2 χ^2	973.50***	1567.75***	621.96***	1072.88***

This Table shows the Heckman second stage estimation by OLS (Heckman's two-step estimator) of Equation (11). The selection equation (Column 1 in Table 4) models firms' propensity to diversify. *ExcessValue* is regressed on *DIVPAT(THEO)*, which is the diversification pattern index based on the Euclidean distance and with the extreme AiPD pattern of reference theoretically specified. Firm size (*LTA*) and its squared term (*LTA2*), financial leverage (*LDTA*), profitability (*EBITsales*), level of investment (*CAPEXsales*), industry effect (*industry dummies*), and time effect (*year dummies*) are controlled in the estimations. The Inverse Mills Ratio ((λ_i) is included as an additional regressor to correct potential self-selection bias in the sample. The Wald test contrasts the null hypothesis of no joint significance of the explanatory variables. Standard error is shown in parentheses under coefficients. ****, ** and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7
Summary statistics of two-dimensional index component variables for the calibration sample (1998-2010)

| | Calibration sample defined by Q (dumQ_BOT=1) | | | | Calibration sample defined by RDsales (dumRDsales_BOT=1) | | | |
	N	Mean	Median	STD	N	Mean	Median	STD
INTER	1660	0.1426	0	0.2827	3433	0.1322	0	0.2872
INTRA	1660	0.5945	1	0.4596	3433	0.5868	1	0.4649

This Table displays the summary statistics of the two-dimensional index component variables (*INTER* and *INTRA*). *INTER* denotes the Gini Index; and *INTRA* is the ratio of a firm's segments displaying commitment ratios above or equal to 1, over the total number of a firm's segments. *dumQ_BOT* is a dummy variable which equals 1 if the firm belongs to the bottom ten percent of sample firms according to the variable *Q* (Cao *et al*. 2008), and zero otherwise. *dumRDsales_BOT* is a dummy variable which equals 1 if the firm belongs to the bottom ten percent of sample firms according to the variable *RDsales* (Mehran, 1995), and zero otherwise.

Table 8
Diversification pattern index (based on calibration samples) and Excess Value [Heckman second stage - Eq. (11)]

	Model ExcessValue = f(DIVPAT(CAL), control variables)							
	Dependent variable: Excess Value				Dependent variable: Excess Value (without extremes)			
	Index based on calibration sample of Q	Index based on calibration sample of Q with industry and time dummies	Index based on calibration sample of RDsales	Index based on calibration sample of RDsales with industry and time dummies	Index based on calibration sample of Q	Index based on calibration sample of Q with industry and time dummies	Index based on calibration sample of RDsales	Index based on calibration sample of RDsales with industry and time dummies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
constant	-1.6840***	-1.8309***	-1.9575***	-1.9357***	-0.9779***	-1.1923***	-1.1977***	-1.2171***
	(0.1330)	(0.1424)	(0.1444)	(0.1575)	(0.1209)	(0.1280)	(0.1312)	(0.1414)
DIVPAT(CAL_Q)	0.4790***	0.7130***			0.3894***	0.5844***		
	(0.0445)	(0.0476)			(0.0396)	(0.0426)		
DIVPAT(CAL_RD)			0.4995***	0.6856***			0.3563***	0.4969***
			(0.0453)	(0.0510)			(0.0397)	(0.0454)
control variables								
LTA	0.3089***	0.2398***	0.3778***	0.3022***	0.1306***	0.0828**	0.1844***	0.1166***
	(0.0409)	(0.0392)	(0.0444)	(0.0438)	(0.0371)	(0.0356)	(0.0404)	(0.0395)
LDTA	-0.0389	-0.0204	-0.1341	-0.0595	-0.0583	-0.1011	-0.1353*	-0.1324*
	(0.0776)	(0.0761)	(0.0869)	(0.0873)	(0.0679)	(0.0668)	(0.0744)	(0.0753)
EBITsales	0.3736***	0.4761***	0.8921***	0.9641***	0.3221***	0.4093***	0.6903***	0.7420***
	(0.0766)	(0.0727)	(0.0899)	(0.0873)	(0.0678)	(0.0647)	(0.0779)	(0.0762)
CAPEXsales	0.6788***	1.4211***	0.8502***	1.4579***	0.4463***	1.0236***	0.5724***	0.9592***
	(0.1258)	(0.1427)	(0.1390)	(0.1602)	(0.1081)	(0.1236)	(0.1168)	(0.1366)
LTA2	-0.0175***	-0.0123***	-0.0221***	-0.0165***	-0.0054*	-0.0014	-0.0089***	-0.0034
	(0.0032)	(0.0031)	(0.0035)	(0.0034)	(0.0029)	(0.0028)	(0.0031)	(0.0031)
Inverse Mills Ratio (λ_i)	-0.1247***	-0.1221***	-0.1433***	-0.1776***	-0.0822***	-0.0821***	-0.0344	-0.0602*
	(0.0298)	(0.0316)	(0.0334)	(0.0372)	(0.0258)	(0.0275)	(0.0282)	(0.0318)
Industry dummies	NO	YES	NO	YES	NO	YES	NO	YES
Year dummies	NO	YES	NO	YES	NO	YES	NO	YES
No. of Obs.	13234	13234	12035	12035	12373	12373	11009	11009
No. Censored Obs.	9537	9537	8449	8449	8902	8902	7737	7737
No. Uncensored Obs.	3697	3697	3586	3586	3471	3471	3272	3272
Wald Chi2 χ^2	567.28***	1308.16***	762.87***	1313.54***	370.06***	1038.95***	479.43***	926.28***

This Table shows the Heckman second stage estimation by OLS (Heckman's two-step estimator) of Equation (11). The selection equation (Column 1 in Table 4) models firms' propensity to diversify. *ExcessValue* is regressed on the alternative specifications of the diversification pattern index based on calibration samples. These alternative specifications of the index are denoted by DIVPAT(CAL_Q) for which the calibration sample is based on *Q*, and DIVPAT(CAL_RD) for which the calibration sample is based on RDsales. Firm size (*LTA*) and its squared term (*LTA2*), financial leverage (*LDTA*), profitability (*EBITsales*), level of investment (*CAPEXsales*), industry effect (*industry dummies*), and time effect (*year dummies*) are controlled in the estimations. The Inverse Mills Ratio ((λ_i) is included as an additional regressor to correct

potential self-selection bias in the sample. The Wald test contrasts the null hypothesis of no joint significance of the explanatory variables. Standard error is shown in parentheses under coefficients. ****, ** and * denote statistical significance at the 1%, 5%, and 10% level, respectively

Table 9
Diversification pattern index and Excess Value [Eq. (11)] – Diversified firms subsample OLS

	Model ExcessValue = f(DIVPAT, control variables)											
	Dependent variable: Excess Value						Dependent variable: Excess Value (without extremes)					
	Baseline index	Baseline index with industry and time dummies	Index based on calibration sample of Q	Index based on calibration sample of Q with industry and time dummies	Index based on calibration sample of RDsales	Index based on calibration sample of RDsales with industry and time dummies	Baseline index	Baseline index with industry and time dummies	Index based on calibration sample of Q	Index based on calibration sample of Q with industry and time dummies	Index based on calibration sample of RDsales	Index based on calibration sample of RDsales with industry and time dummies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
constant	-2.3113*** (0.1235)	-2.3997*** (0.1339)	-1.8617*** (0.1265)	-2.0157*** (0.1356)	-2.1733*** (0.1358)	-2.2179*** (0.1479)	-1.3285*** (0.1115)	-1.4385*** (0.1201)	-1.0831*** (0.1166)	-1.3075*** (0.1234)	-1.2476*** (0.1249)	-1.3083*** (0.1345)
DIVPAT(THEO)	0.6779*** (0.0448)	0.7396*** (0.0466)					0.4852*** (0.0390)	0.5161*** (0.0410)				
DIVPAT(CAL_Q)			0.4258*** (0.0428)	0.6776*** (0.0473)					0.3503*** (0.0377)	0.5574*** (0.0421)		
DIVPAT(CAL_RD)					0.4572*** (0.0444)	0.6399*** (0.0507)					0.3451*** (0.0387)	0.4794*** (0.0449)
control variables												
LTA	0.3832*** (0.0397)	0.3350*** (0.0394)	0.3301*** (0.0407)	0.2568*** (0.0394)	0.4028*** (0.0442)	0.3299*** (0.0440)	0.1678*** (0.0352)	0.1392*** (0.0350)	0.1417*** (0.0371)	0.0920*** (0.0357)	0.1898*** (0.0399)	0.1246*** (0.0397)
LDTA	-0.1090 (0.0773)	-0.1124 (0.0789)	0.0009 (0.0773)	-0.0042 (0.0769)	-0.0851 (0.0864)	-0.0311 (0.0883)	-0.1021 (0.0655)	-0.1472** (0.0675)	-0.0313 (0.0676)	-0.0895 (0.0675)	-0.1234* (0.0738)	-0.1220 (0.0759)
EBITsales	0.6158*** (0.0729)	0.6824*** (0.0718)	0.3787*** (0.0766)	0.4692*** (0.0733)	0.8678*** (0.0897)	0.9266*** (0.0875)	0.5491*** (0.0633)	0.5974*** (0.0631)	0.3274*** (0.0679)	0.4053*** (0.0653)	0.6830*** (0.0777)	0.7263*** (0.0765)
CAPEXsales	0.7187*** (0.1234)	1.4143*** (0.1429)	0.5890*** (0.1240)	1.3033*** (0.1408)	0.7393*** (0.1366)	1.2952*** (0.1580)	0.4392*** (0.1028)	0.9111*** (0.1214)	0.3896*** (0.1067)	0.9480*** (0.1223)	0.5471*** (0.1150)	0.9071*** (0.1353)
LTA2	-0.0217*** (0.0031)	-0.0175*** (0.0031)	-0.0183*** (0.0032)	-0.0129*** (0.0031)	-0.0231*** (0.0035)	-0.0175*** (0.0035)	-0.0072*** (0.0028)	-0.0041 (0.0027)	-0.0057** (0.0029)	-0.0016 (0.0028)	-0.0091*** (0.0031)	-0.0036 (0.0031)
Industry dummies	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Year dummies	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
No. of Obs.	4507	4507	3697	3697	3586	3586	4097	4097	3471	3471	3272	3272
F-test	166.36***	22.57***	100.99***	19.05***	134.67***	19.12***	105.03***	15.25***	63.69***	14.78***	81.34***	13.07***
Adjusted-R^2	0.1805	0.2510	0.1396	0.2547	0.1828	0.2613	0.1322	0.1958	0.0978	0.2175	0.1284	0.2053

This Table shows the OLS estimations of Equation (11) in the subsample of diversified firms. *ExcessValue* is regressed on the diversification pattern index. *DIVPAT(THEO)*, *DIVPAT(CAL_Q)* and *DIVPAT(CAL_RD)*, denote the alternative specifications for the diversification pattern index based on the Euclidean distance. Firm size (*LTA*) and its squared term (*LTA2*), financial leverage (*LDTA*), profitability (*EBITsales*), level of investment (*CAPEXsales*), industry effect (*industry dummies*), and time effect (*year dummies*) are controlled in the estimations. The F- test contrasts the null hypothesis of no joint significance of the explanatory variables. Standard error is shown in parentheses under coefficients. ****, ** and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 10
Diversification pattern index and Excess Value [Heckman second stage - Eq. (12)]]
Additional sensitivity tests adding the dimension scope

	Model ExcessValue = f(DIVPAT, DIVER, control variables)							
	Dependent variable: Excess Value				Dependent variable: Excess Value (without extremes)			
	Only control variables	Full model with the baseline index	Full model with index based on calibration sample of Q	Full model with index based on calibration sample of RDsales		Only control variables	Full model with the baseline index	Full model with index based on calibration sample of Q	Full model with index based on calibration sample of RDsales
	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)
constant	-1.6737*** (0.1433)	-2.2248*** (0.1437)	-1.9293*** (0.1454)	-2.0420*** (0.1612)		-0.9170*** (0.1252)	-1.3141*** (0.1276)	-1.2282*** (0.1312)	-1.2521*** (0.1449)
DIVPAT(THEO)		0.8517*** (0.0533)					0.5548*** (0.0468)		
DIVPAT(CAL_Q)			0.8084*** (0.0561)					0.6170*** (0.0500)	
DIVPAT(CAL_RD)				0.7790*** (0.0595)					0.5265*** (0.0527)
Degree of diversification (DIVER)									
MHERF	-0.3590*** (0.0662)	0.1723** (0.0724)	0.2401*** (0.0746)	0.2544*** (0.0838)		-0.3180*** (0.0567)	0.0151 (0.0624)	0.0816 (0.0656)	0.0805 (0.0730)
control variables									
LTA	0.3530*** (0.0402)	0.3066*** (0.0393)	0.2300*** (0.0393)	0.2938*** (0.0438)		0.1462*** (0.0354)	0.1279*** (0.0349)	0.0806** (0.0355)	0.1146*** (0.0395)
LDTA	-0.1299 (0.0804)	-0.1326* (0.0782)	-0.0148 (0.0760)	-0.0490 (0.0873)		-0.1592** (0.0681)	-0.1614** (0.0670)	-0.0986 (0.0668)	-0.1284* (0.0753)
EBITsales	0.6809*** (0.0733)	0.7056*** (0.0716)	0.4902*** (0.0728)	0.9804*** (0.0873)		0.5803*** (0.0637)	0.6044*** (0.0628)	0.4138*** (0.0648)	0.7475*** (0.0764)
CAPEXsales	1.7518*** (0.1471)	1.5499*** (0.1441)	1.4261*** (0.1426)	1.4603*** (0.1601)		1.0960*** (0.1238)	0.9828*** (0.12222)	1.0268*** (0.1236)	0.9615*** (0.1366)
LTA2	-0.0176*** (0.0032)	-0.0166*** (0.0031)	-0.0119*** (0.0031)	-0.0163*** (0.0034)		-0.0036 (0.0028)	-0.0039 (0.0027)	-0.0014 (0.0028)	-0.0033 (0.0031)
Inverse Mills Ratio (λ_i)	-0.0487 (0.0318)	-0.1647*** (0.0319)	-0.1306*** (0.0317)	-0.1840*** (0.0372)		-0.0119 (0.0267)	-0.0906*** (0.0271)	-0.0853*** (0.0276)	-0.0625** (0.0319)
Industry dummies	YES	YES	YES	YES		YES	YES	YES	YES
Year dummies	YES	YES	YES	YES		YES	YES	YES	YES
No. of Obs.	16554	16554	13234	12035		15104	15104	12373	11009
No. Censored Obs.	12047	12047	9537	8449		11007	11007	8902	7737
No. Uncensored Obs.	4507	4507	3697	3586		4097	4097	3471	3272
Wald Chi2 χ^2	1256.87***	1574.83***	1321.49***	1325.31***		904.47***	1072.92***	1040.78***	927.73***

This Table shows the Heckman second stage estimation by OLS (Heckman's two-step estimator) of eq. (12). The selection equation (Column 1 in Table 4) models firms' propensity to diversify. *ExcessValue* is regressed on the diversification pattern index and the degree of diversification. *DIVPAT(THEO)*, *DIVPAT(CAL_Q)* and *DIVPAT(CAL_RD)*, denote the alternative specifications for the diversification pattern index based on the Euclidean distance. *MHERF* represents the Herfindahl index and

measures the degree of diversification. Firm size (*LTA*) and its squared term (*LTA2*), financial leverage (*LDTA*), profitability (*EBITsales*), level of investment (*CAPEXsales*), industry effect (*industry dummies*), and time effect (*year dummies*) are controlled in the estimations. The Inverse Mills Ratio ((λ_i) is included as an additional regressor to correct potential self-selection bias in the sample. The Wald test contrasts the null hypothesis of no joint significance of the explanatory variables. Standard error is shown in parentheses under coefficients. ****, ** and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 11
Diversification pattern index and Excess Value controlling for diversification and growth opportunities
[Heckman second stage - Eq. (13)]

	Model ExcessValue = f(DIVPAT, DIVER, GROWTH, control variables)					
	Dependent variable: Excess Value					
	Full model with the baseline index and *lnAge*	Full model with index based on calibration sample of Q and l*nAge*	Full model with index based on calibration sample of RDsales and l*nAge*	Full model with the baseline index and RDA	Full model with index based on calibration sample of Q and RDA	Full model with index based on calibration sample of RDsales and RDA
	(1)	(2)	(3)	(4)	(5)	(6)
constant	-3.1943***	-1.9021***	-2.7967***	-2.1835***	-1.7540***	-1.9799***
	(0.6205)	(0.7038)	(0.6728)	(0.2148)	(0.2183)	(0.2667)
DIVPAT(THEO)	1.1522***			0.9293***		
	(0.1952)			(0.0771)		
DIVPAT(CAL_Q)		1.0755***			0.9259***	
		(0.2134)			(0.0824)	
DIVPAT(CAL_RD)			0.6801***			0.9536***
			(0.1913)			(0.0911)
Degree of diversification (DIVER)						
MHERF	0.0654	-0.1837	-0.7558***	0.2561**	0.4639***	0.4755***
	(0.2658)	(0.3143)	(0.2806)	(0.1035)	(0.1085)	(0.1303)
Growth opportunities						
lnAge	-0.0075	-0.0013	0.0306			
	(0.0603)	(0.0604)	(0.0581)			
RDA				1.8372***	1.1221***	0.9028**
				(0.3029)	(0.3098)	(0.4017)
control variables						
LTA	0.8005***	0.4968***	0.7490***	0.2630***	0.1338**	0.2067***
	(0.1381)	(0.1575)	(0.1449)	(0.0570)	(0.0580)	(0.0703)
LDTA	-0.5736**	-0.3097	-0.5027*	-0.0543	-0.1091	0.0496
	(0.2869)	(0.2971)	(0.2891)	(0.1221)	(0.1206)	(0.1529)
EBITsales	0.5319**	0.0662	0.7007***	1.0252***	0.8006***	1.3718***
	(0.2269)	(0.2684)	(0.2644)	(0.1120)	(0.1135)	(0.1477)
CAPEXsales	1.7153***	2.9179***	1.9407***	1.8040***	1.5405***	2.2927***
	(0.4328)	(0.5579)	(0.5335)	(0.3184)	(0.3151)	(0.4723)
LTA2	-0.0550***	-0.0302**	-0.0497***	-0.0136***	-0.0054***	-0.0115**
	(0.0114)	(0.0131)	(0.0121)	(0.0044)	(0.0045)	(0.0054)
Inverse Mills Ratio (λ_i)	-0.3918***	-0.5756***	-0.3008*	-0.1875***	-0.1202***	-0.1661***
	(0.1449)	(0.1549)	(0.1677)	(0.0464)	(0.0466)	(0.0597)
Industry dummies	YES	YES	YES	YES	YES	YES
Year dummies	YES	YES	YES	YES	YES	YES
No. of Obs.	12469	9840	8791	14127	11269	9899
No. Censored Obs.	12047	9537	8449	12047	9537	8449
No. Uncensored Obs.	422	303	342	2080	1732	1450
Wald Chi2 χ^2	460.43***	283.45***	479.00***	848.99***	656.81***	658.66***

This Table shows the Heckman second stage estimation by OLS (Heckman's two-step estimator) of Equation (13). The selection equation (Column 1 in Table 4) models firms' propensity to diversify. *ExcessValue* is regressed on the diversification pattern index, degree of diversification and growth opportunities. *DIVPAT(THEO)*, *DIVPAT(CAL_Q)* and *DIVPAT(CAL_RD)*, denote the alternative specifications for the diversification pattern index based on the Euclidean distance. *MHERF* represents the Herfindahl index and measures the degree of diversification. Growth opportunities are proxied by either *lnAge* (natural logarithm of firm age) or *RDA* (the ratio of annual R&D expenditures and beginning-of-year book assets) alternatively. Firm size (*LTA*) and its squared term (*LTA2*), financial leverage (*LDTA*), profitability (*EBITsales*), level of investment (*CAPEXsales*), industry effect (*industry dummies*), and time effect (*year dummies*) are controlled in all estimations. The Inverse Mills Ratio ((λ_i) is included as an additional regressor to correct potential self-selection bias in the sample. The Wald test contrasts the null hypothesis of no joint significance of the explanatory variables. Standard error is shown in parentheses under coefficients. ****, ** and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 12
Industry pre-emption heterogeneity, Diversification pattern index and Excess Value
[Heckman second stage - Eq. (12)]

	INDUSTRY PREEMPTION HETEROGENEITY: Subsamples of firm-year observations by risk of pre-emption in the core industry		
	DIVPAT(THEO)	DIVPAT(CAL_Q)	DIVPAT(CAL_RD)
1st tercile *PREEMPT*	0.8341***	0.8561***	0.7007***
	(0.0900)	(0.0953)	(0.0938)
2nd tercile *PREEMPT*	0.6248***	0.5607***	0.5864***
	(0.0767)	(0.0805)	(0.0796)
3rd tercile *PREEMPT*[a]	0.4709***	0.4622***	0.4334***
	(0.0228)	(0.0246)	(0.0254)

This Table examines industry heterogeneity based on *PREEMPT*. *PREEMPT* measures risk of pre-emption, calculated as the natural logarithm of the number of firms operating in the same 2-digit SIC code industry as the core business of the corresponding firm. This Table shows the coefficients of the alternative specifications for the diversification pattern index of the Heckman second stage estimations (Heckman's two-step estimator) of equation (12). Regressions are performed separately for three different subsamples formed by terciles of *PREEMPT* levels. The selection equation (Column 1 in Table 4) models firms' propensity to diversify. *ExcessValue* is regressed on the diversification pattern index and the degree of diversification. *DIVPAT(THEO)*, *DIVPAT(CAL_Q)* and *DIVPAT(CAL_RD)*, denote the alternative specifications for the diversification pattern index based on the Euclidean distance. *MHERF* represents the Herfindahl index and measures the degree of diversification. Firm size (*LTA*) and its squared term (*LTA2*), financial leverage (*LDTA*), profitability (*EBITsales*), level of investment (*CAPEXsales*), industry effect (*industry dummies*), and time effect (*year dummies*) are controlled in all estimations. The Inverse Mills Ratio ((λ_i) is included as an additional regressor to correct potential self-selection bias in the sample. The Wald test contrasts the null hypothesis of no joint significance of the explanatory variables. Standard error is shown in parentheses under coefficients. ***, ** and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

[a] Regressions in the third tercile of the *PREEMPT* subsample are performed by OLS due to the lack of statistical significance of the Inverse Mills Ratio (λ_i) in the corresponding Heckman regressions.

Table 13
Dynamic models of Diversification pattern index and Excess Value controlling for diversification
[Heckman second stage - Eq. (14)]

Model ExcessValue = f(DIVPAT(THEO), DIVPAT(-1), DIVPAT(-2), DIVER, control variables)		
	Dependent variable: Excess Value	
	With one lag of the baseline index (1)	With two lags of the baseline index (2)
constant	-2.0479*** (0.1695)	-1.8043*** (0.2033)
DIVPAT(THEO)	0.7389*** (0.0665)	0.6716*** (0.0791)
DIVPAT(-1)	0.1935*** (0.0471)	0.2080*** (0.0652)
DIVPAT(-2)		0.0636 (0.0527)
Degree of diversification (DIVER)		
MHERF	0.2620*** (0.0822)	0.3764*** (0.0951)
control variables		
LTA	0.2359*** (0.0460)	0.2001*** (0.0548)
LDTA	-0.0888 (0.0909)	0.0352 (0.1077)
EBITsales	0.9758*** (0.0927)	1.1752*** (0.1148)
CAPEXsales	1.3955*** (0.1851)	1.3318*** (0.2276)
LTA2	-0.0120*** (0.0036)	-0.0104** (0.0042)
Inverse Mills Ratio (λ_i)	-0.1556*** (0.0344)	-0.1945*** (0.0396)
Industry dummies	YES	YES
Year dummies	YES	YES
No. of Obs.	15211	14351
No. Censored Obs.	12047	12047
No. Uncensored Obs.	3164	2304
Wald Chi2 χ^2	1195.22***	900.51***

This Table shows the Heckman second stage estimation by OLS (Heckman's two-step estimator) of Equation (13). The selection equation (Column 1 in Table 4) models firms' propensity to diversify. *ExcessValue* is regressed on the diversification pattern index (*DIVPAT(THEO)*), its lags, and the degree of diversification. *DIVPAT(-1)* is the one-period lagged *DIVPAT(THEO)*, and *DIVPAT(-2)* is the two-period lagged *DIVPAT(THEO)*. *MHERF* represents the Herfindahl index and measures the degree of diversification. Firm size (*LTA*) and its squared term (*LTA2*), financial leverage (*LDTA*), profitability (*EBITsales*), level of investment (*CAPEXsales*), industry effect (*industry dummies*), and time effect (*year dummies*) are controlled in all estimations. The Inverse Mills Ratio ((λ_i) is included as an additional regressor to correct potential self-selection bias in the sample. The Wald test contrasts the null hypothesis of no joint significance of the explanatory variables. Standard error is shown in parentheses under coefficients. ****, ** and * denote statistical significance at the 1%, 5%, and 10% level, respectively.