Apples to Apples:
The Economic Benefit of Corporate Diversification

John E. Hund, Donald Monk, and Sheri Tice*

March 15, 2012

Abstract

Valuation differences between focused and diversified firms are misleading due to comparing relatively small, young, and more volatile focused firms with larger, older and less volatile diversified firms. The largest diversified firms are also the most valuable and as a consequence the value-weighted diversification "discount" is actually a premium. We highlight this issue by showing that diversified firms have a value-weighted average economy-wide gain of $885 billion annually relative to imputed firm values based on focused firms. Diversified firms also comprise 75% on average of the market value of the S&P 500, so among large firms, they are considered more valuable. We also test whether the diversification "discount" is an artifact of comparing focused and diversified firms which differ substantially on dimensions related to the uncertainty of their growth rates. After accounting for firm differences using both propensity score methods and coarsened exact matching, we show that diversified firms trade at a premium compared to similar focused firms. Our findings reveal higher values for diversified firms relative to comparable focused firms---a result that is inconsistent with prior theoretical research concluding that diversification destroys value.

JEL classification: G32, G34
Keywords: Organizational Structure, Diversification, Firm Valuation, Coarsened Exact Matching

*John Hund (jhund@rice.edu) is at the Jones Graduate School of Business at Rice University, Donald Monk (dmonk@business.rutgers.edu) is at Rutgers Business School at Rutgers University and Sheri Tice (stice@tulane.edu) is at the A.B. Freeman School of Business at Tulane University. We appreciate comments received from brown bag participants at the University of Texas at Austin and Rutgers University.

1. Introduction

Diversified firms are older and have higher sales and market values on average than focused firms. Yet diversified firms have a lower market value than the market value of a hypothetical firm constructed from the market-sales ratios of focused firms, a fact which has been broadly labeled the "diversification discount" in the voluminous literature which followed Berger and Ofek (1995). This fact has led researchers to conclude that firm diversification destroys value both at the aggregate (economy-wide) and at the firm level. We show in this paper that neither of these conclusions is true.

Our first tests focus on the overall value of diversified firms within the realm of public firms using the conventional methodology to calculate discounts. We show that the largest diversified firms tend to trade at a premium and not a discount. Though the equal-weighted diversification discount is negative, the value-weighted discount is positive. Moreover, the value-weighted discount is positive in every year except two from 1977-2009, resulting in an average gain to the economy of 885 billion dollars a year, measured in nominal dollars. Diversified firms also comprise 75% on average of the market value of the S&P 500, indicating that among large firms, they are considered more valuable in that they are consistently chosen to be included in this well respected index.

Next, we turn our attention to the weaknesses of the current methodology for comparing diversified firm values to focused firm values. Typically, Compustat SIC codes are used as the first and only categorization of firms before imputed values are created using within industry market-to-sales ratios.[1] We show that industry market-sales ratios are extremely unstable across industries and time, even for industries defined by adjacent SIC codes, and often segment and firm SIC codes are classified inconsistently in Compustat data. Going further, we show that the diversification discount at the firm level is an artifact of matching older, larger and less volatile firms with younger, smaller and more volatile firms. When we match firms on age, sales, and

[1] Villalonga (2004b) is a notable exception. She finds that the use of industry data other than SIC codes taken from Compustat segment data results in a diversification premium.

return volatility within year and 1-digit SIC code, diversified firms have higher market values than their matched focused peers using propensity matching and statistically equal market values using coarsened exact matching. We conclude that the diversification discount is an artifact of comparing firms with different valuation-relevant characteristics, and diversified firms do not to destroy value when compared properly.

Our paper revisits the major themes that have developed in the diversification literature since the seminal research of Berger and Ofek (1995). Much of the literature attempts to explain the reason that diversified firms trade at a discount to their imputed values. Papers in this vein of the literature argue that the complex diversified form leads to inefficient investment (Shin and Stulz (1998), Rajan, Servaes, and Zingales (2000), Scharfstein and Stein (2000), Lamont and Polk (2002), Ozbas and Scharfstein (2010)), that diversified firms tend to acquire poorly run firms (Graham, Lemmon, and Wolf (2002)), or that diversified firms have poor governance (Hoechle, Schmid, Walter, and Yermack (2011), Denis, Denis and Sarin (1997)). We show that diversified firms trade at a premium when we match apples to apples, create a large amount of wealth annually, and consistently represent approximately 75% of the market value of the S&P 500. As such, our findings cast doubt on the interpretation of papers in the literature that argue that diversified firms are poorly run and destroy value.

Our paper is one of a few that shows diversified firms appear to be well run firms that create value. Our findings are consistent with studies, such as Khanna and Tice (2001), Maksimovic and Phillips (2002), and Tate and Yang (2011), that find that diversified firms appear to be deploying resources efficiently. Like Campa and Kedia (2002), Maksimovic and Phillips (2002), Bevelander (2002), Villalonga (2004a), Borghesi, Houston, and Naranjo (2007), and Colak (2010) we argue that diversified firms are different than focused firms. However, we are the first to point out the aggregate positive effects of diversification in the economy. We are also the first to show that the concentration of large firms among diversified firms trading at a premium to

focused firms converts the diversification discount to a premium when firms are value-weighted. Our work also builds on the assertions in Hund, Monk, and Tice (2010) who argue that the diversification discount results from differences in the uncertainty of growth rates between focused and diversified firms. We are the first to show that the entire diversification discount disappears and converts to a premium in a comprehensive universe of all firms from 1977-2009 once diversified and focused firms are correctly compared after controlling for variables associated with the uncertainty of growth rates.

The paper proceeds as follows. Section 2 describes our data and Section 3 summarizes key statistics documenting differences between focused and diversified firms. Section 4 contains results on the economic benefits of diversification to the aggregate economy and shows the value-weighted diversification premium. Section 5 describes problems with matching firms by industry. Section 6 presents matching estimates on the diversification discount using both propensity score and coarsened exact matching methods, and Section 7 concludes.

2. Data and Sample Construction

2.1 Data Sources and Diversified Indicator

Our empirical analysis begins with merged data from the segment- and firm-level Compustat Industrial Annual files for the period 1977-2009 and return and market value data from the CRSP monthly returns files. Firm annual return, volatility, and market value variables are calculated at the fiscal year end dates from Compustat using the monthly CRSP stock return data. S&P 500 membership data are extracted from the Compustat Index Constituents file. Firm-years are dropped from the sample according to the Berger and Ofek (1995) requirements that firms have no segments in the financial services industry (SIC 6000-6999), total firm sales are above $20 million, and aggregated firm segment sales are within 1% of firm-level data. We also remove regulated utilities (SIC 4900-4941) and firms that do not report sales and four-digit SIC

codes for all of their segments.[2] To address the complexities introduced by the new segment reporting rule SFAS 131 (which went into effect in 1998) and to deal with the problem of pseudo-conglomerates (as in Sanzhar, (2006)), we perform the aggregation procedure detailed in Hund, Monk, & Tice (2010).[3] Diversified firms are identified as firms with more than one segment following the segment aggregation procedure.

2.2 Variable definitions

The diversification premium or discount is calculated at the firm level by following the procedures in Berger & Ofek (1995). We calculate excess value (EV) as the log-ratio of total capital to the imputed value for the firm. The imputed value for the firm is calculated by multiplying the median ratio of total capital to sales for focused firms in a segment's industry by the segment's reported sales and then summing over the number of segments in the firm. Specifically, excess value is:

$$I(V) = \sum_{i=1}^{n} S_i * \left(Ind_i \left(\frac{V}{S} \right)_{mf} \right),$$

$$EV = ln\big(V/I(V)\big),$$

where *I(V)* is the imputed value, *V* is the firm total capital (market value of equity plus book value of debt), S_i is sales reported for segment *i*, $Ind_i\ (V/S)_{mf}$ is the ratio of total capital to sales for the median focused firm in the same industry as segment *i*, and *n* is the number of segments in the firm.[4] The matched segment median value comes from the finest SIC code level (two-,

[2] As a robustness check, we include financial firms (except the government sponsored entities Fannie Mae and Freddie Mac). Our findings throughout the paper remain the same using this extended sample, though for brevity we report only one summary line of results in Table 3.

[3] Effectively this procedure adds sales from multiple segments reported in the same four-digit SIC code into one segment with that four-digit SIC code, and then re-classifies as focused those firms whose segments were all within the same four-digit SIC code. For more details, see Berger and Hann (2003).

[4] For the results presented in the main section of the paper, we use the market value of equity computed from CRSP at the end of the firm's fiscal year. We have reproduced all of our results using market values computed from

three-, or four-digit) with at least five focused firms. Firms with excess values greater than or equal to zero are designated as "premium" firms, and firms with excess values less than zero are designated as "discount" firms.

Two important points about the excess value measure should be emphasized. First, the excess value measure is the log-ratio of the firm's market-to-sales ratio over the sales-weighted average market-to-sales ratio of the median focused firm in each industry in which the firm operates. The diversification discount is essentially a statement that on average, diversified firms have lower market-to-sales ratios than the median focused firms that operate in their industries. If market-to-sales ratios vary consistently along other dimensions than industry (as we document for size and age), then the diversification discount devolves into a statement about those confounding dimensions rather than one about corporate form. Second, we also note that the excess value measure defines a distribution of excess values for **both** focused and diversified firms; only the median focused firm will by definition have an excess value of zero. There are substantial numbers of "discount" and "premium" focused firms, just as there are substantial numbers of "discount" and "premium" diversified firms.

We carefully consider the construction of firm age, since it is such a critical variable in the assessment of the growth rate uncertainty of the firm. The age of the firm at IPO has fallen substantially during our sample, so simply using the first appearance in the database or the listing date severely understates the age of older firms in the sample and exacerbates the bias in comparing older firms' market-to-sales ratios with younger. We define firm age using data containing firm "birth" dates from Jovanovic and Rousseau (2001) that has been supplemented

---

Compustat prices and shares outstanding and using average monthly market values for the 12 months in the firm's fiscal year.

by Fink, Fink, Grullon, and Weston (2011). For the few firms remaining without birth dates in these databases, we calculate birth dates using the first listing date in Compustat.

Our other variables are standard accounting and returns measures. *Profit Margin* is defined as EBITDA over sales, *Earnings* are defined as income before extraordinary items plus deferred taxes and tax credits, and *Sales* are total firm revenues. *Return Volatility* is the standard deviation of the monthly stock returns during the firm's fiscal year, and *Excess Return Volatility* is the standard deviation of the monthly excess stock returns during the firm's fiscal year where excess stock returns are defined as the difference between the firm's stock return and the value-weighted market return for that month.

3. Summary Statistics

Table 1 shows the dramatic difference between focused and diversified firms and between the discount and premium sub-categories within these designations. Diversified firms have over twice the market value, assets and sales of focused firms and are nearly twice as old. Diversified firms are more profitable and have significantly lower return volatility as should be expected from more mature, larger firms. Median values are substantially lower than the means for the market values and accounting measures, pointing out the severe skewness in these distributions that we discuss in the following section.

Across both diversified and focused firms, premium firms have much larger market values than discount firms, and in particular, diversified premium firms have over 3.5 times the market value of diversified discount firms. An immediate consequence of this is that the value-weighted excess value (EV) for diversified firms exceeds the equal-weighted one, and in subsequent sections we show that this is true and the difference is substantial. Premium diversified firms have much higher sales and earnings than discount ones, whereas premium focused firms are

more profitable, but have similar sales compared to discount focused firms. Interestingly, there is very little variation in age across discount and premium firms either for focused or diversified firms. However, diversified firms are clearly older than focused firms.

Diversified firms are older, much larger on all dimensions and more profitable than focused firms, consistent with the idea that diversified firms are very different than focused firms. Diversified firms also have lower return and excess return volatility. Taken together, these facts support the hypothesis that diversified firms have lower uncertainty about their growth rates, and potentially, lower market-to-sales ratios than focused firms for reasons that are entirely consistent with value maximizing behavior in an older, more mature firm. The over ten thousand premium diversified firm-years in our sample also represent the oldest, highest sales, most profitable, and biggest market value firms in our sample.[5]

4. Economic Significance of Diversified Firms

4.1 Decile Ranks and Excess Values

In Table 2 we more formally examine the skewness in the distributions of market values, sales, and age and the interaction of these values with the excess value measure. We calculate decile breakpoints within each group of focused and diversified firms (note that this means that the breakpoints in each decile will differ substantially), and then calculate the mean and median excess values, market value of equity (Panel A), sales (Panel B), and age (Panel C). One important conclusion to be drawn from this table is that the mean market value is dramatically increasing across deciles for both focused and diversified firms, but much more rapidly for diversified than for focused firms. Whereas the left tails of both distributions begin at

[5] We show later in the paper that these firms dominate the S&P 500, a fact consistent with the hypothesis that diversification is not necessarily a value destroying firm form.

approximately the same place, the highest decile of diversified firms have an average market value that is nearly one and a half times the market value of the highest decile of focused firms. The excess value measure is also increasing monotonically across deciles, and reaches a premium in decile 8 of the diversified distribution. The fact that the largest market value firms are on average premium firms suggests that diversification adds to aggregate economic value, not subtracts from it as argued by the previous literature, most notably Berger and Ofek (1995). We test this assertion more explicitly in the next section.

Panels B and C of Table 2 show that sales and age deciles have similarly skewed mean values as market values, but a relationship with excess value is not as evident. Mean and median excess values clearly increase with sales, but there is not a monotonic relationship as shown in Panel A. The relationship between age deciles and excess value measures is even weaker as shown in Panel B, although importantly, older focused firms seem to exhibit a discount just as older diversified firms do, supporting the idea that it is uncertainty that drives the diversification discount. Indeed, if agency costs and value destruction were driving the discount, we would expect to see higher discounts in higher sales deciles (as larger firms should be more prone to empire building and suboptimal cross-subsidization as in Rajan, Servaes, and Zingales (2000) and Scharfstein and Stein, (2000)). Yet, we find exactly the reverse.

We document that size and age are indeed confounding factors in the comparison of diversified and focused firms using the excess value measure. For both diversified and focused firms, we show that larger and older firms have lower market-to-sales ratios than younger and smaller firms, controlling for industry effects. If larger and older firms have less uncertainty about their future growth rates, this implies that the diversification discount can be obtained from a learning model as in Pastor & Veronesi (2003) and Hund, Monk, and Tice (2010).

4.2 Value-weighted vs. Equal-weighted Discounts

Table 3 presents one of the main results of our paper: on average, diversification **creates** value at an aggregate economic level. Because of the skewness in the distribution of market values and the association of higher market values with premiums based on the excess value measure, the value-weighted "discount" is positive. In Table 3 we calculate the average excess value in every year by equal-weighting or value-weighting by market values. The difference is substantial. Whereas the overall discount is -.1082 when the simple average is formed, the overall value-weighted average is a premium of .2395.[6] This difference is driven by the fact that the largest diversified firms are generally premium firms and that discount firms are orders of magnitude smaller. Moreover, diversification provides a **net gain to the economy in every year** from 1977 to 2009 save two, 1983 and 1984, when the discounts are -0.0238 and -0.0054, respectively.

It is incorrect to extrapolate the mean discount to an economic loss to the economy associated with the diversified firm form as has been done in virtually all of the pre-existing literature on the diversification discount. Indeed, Berger and Ofek (1995, pg. 49) note "using the asset multiplier, the mean dollar loss per firm during 1986-91 is $235.1 million, implying a total loss in value for the approximately 850 multi-segment sample firms of $200 billion."[7] In

[6] A mean discount of 11% is virtually identical to that found in most previous studies of the diversification discount, including Hund, Monk, and Tice (2010), Berger and Ofek (1995), and many others.

[7] Even though we aggregate our data to remove duplicate segments and match explicitly to CRSP data, when we restrict our attention to the 1986-1991 period studied by Berger and Ofek (1995), we calculate a mean (median) discount of 9.8% (11.4%) compared to 9.7% (10.6%) in Berger and Ofek (1995) for excess values based on sales multiples. In addition, our discounts are also within 2% of Berger and Ofek (1995) at the 25th and 75th percentiles, and have virtually identical standard deviations. In order to more closely replicate the data in Berger and Ofek (1995), we use a legacy segment file from 1999 and the legacy version of Compustat firm data from 2006 and do not aggregate the data or match to CRSP, and calculate discounts based on both assets and sales multiples. While our sales multiple discounts continue to be virtually identical to those in Berger and Ofek (1995) we calculate a median asset multiple discount of 10.3% vs. their 16.6% discount. Our computed discount is identical to that calculated Campa and Kedia (2002) when they restrict their sample to the 1986-1991 period, and we note as they do, that the difference is likely due to restatements and additions in the Compustat files. Our sample using legacy data has 4464 firms (similar to the 4565 firms in Campa and Kedia (2002)), but significantly more than the 3659 reported by Berger and Ofek (1995). Using our legacy data files, the equal-weighted loss to diversification (using sales multiples

Table 3 we also calculate the aggregate market value of all diversified firms and multiply that value by the percentage gain/loss associated with the equal- or value-weighted excess value measure. Simply averaging the equal-weighted loss across all years leads to an aggregate total "loss" in value of 296 billion dollars per year; however, correctly weighted, diversified firms have an average aggregate market value gain of 885 billion dollars per year relative to their imputed values. When we include financial firms in the sample the disparity between equally weighted and value-weighted loss or gain is even greater as indicated in the last row of Table 3. Far from destroying value, diversification has economic benefits of enormous magnitude. Of course these "benefits" rely on the fact that the discount has been calculated by matching solely on industry and not controlling for confounding effects of size, age and volatility on the market-to-sales ratios of diversified and focused firms. We return to this important point in Section 6, where we show that a slight premium still exists when we control for these differences using parametric and non-parametric matching procedures.[8]

4.3 Results using S&P 500 firms

If diversified firms are in the aggregate adding value to the economy, we would expect to find them disproportionately represented in sets of firms that are selected for quality and importance to the economy. The S&P 500 index is an index of 500 firms selected by committee to be "a leading indicator of U.S. equities, reflecting the risk and return characteristics of the broader large cap universe on an on-going basis." (S&P Factsheet, 2011). The index committee selects stocks that are large, public, financially viable, with excellent liquidity and which, when taken as a whole, represent all sectors of the US economy. The index captures approximately

over the period from 1986-1991) is $196 million per firm and the value-weighted gain to diversification is $100 million per firm.

[8] We also perform the same analysis shown in Table 3 using the sample of focused firms. In unreported results, we find that focused firms have a value-weighted premium, though their premium is less than the one exhibited by diversified firms. Please refer to Section 6 for a more rigorous comparison of diversified and focused firms.

75% of all outstanding market value in the United States equity markets. If diversified firms do indeed add economic benefit to the economy as a whole, then we would expect that the S&P 500 index would be dominated by diversified firms. Conversely, if diversification destroys value, we should expect to see the index of large-cap, profitable, and liquid firms dominated by focused firms. In Table 4 we show that the S&P 500 index is dominated by diversified firms and has been in every year since 1977.[9] Specifically, diversified firms outnumber focused firms in the index by approximately 3:1 and represent on average over 75% of the S&P 500 index market value. These facts are in spite of the focused firms comprising approximately 50% of the opportunity set of firms in the market value range for S&P 500 consideration. It is interesting that focused firm participation in the index peaks in 1996 and 1997, but then it very rapidly declines back to a 3:1 margin of diversified:focused for the 2000-2009 period. It is difficult to reconcile the existence of a value-destructive corporate form with the domination of the S&P 500 index in every year of the past three decades. Rather, diversified firm disproportionate membership in the S&P 500 reflects the fact that diversified firms are older, more profitable, higher market value firms that have lower uncertainty about their growth rates.

5. Problems With Industry Matching In the Excess Value Measure

As noted earlier, the excess value measure of Berger and Ofek (1995) is the log ratio of market-to-sales ratios of the entire firm to the median focused firms matched in the finest possible industry match for each segment. The preference for matching within 4-digit SIC codes creates large discrepancies on dimensions (such as size, volatility, and age) over which market-to-sales ratios are well known to vary predictably. Since these omitted dimensions are embedded non-linearly in the excess value measure, they cannot be controlled for within a linear regression

[9] The total number of firms in the S&P 500 index in Table 4 does not add up to 500 due to absence from the Compustat segment file and the fact that several S&P 500 firms in recent years are not US domiciled. However, even if we allocate all of the "missing" firms to the focused category, diversified firms still represent a vastly disproportionate share of the index.

specification merely by including them as independent variables (as in Borghesi, Houston, and Naranjo (2007)).

In addition, there is ample reason to believe that segment data reported in Compustat is less than perfectly reliable. SFAS 131 (enacted in 1997) gave substantial latitude to corporations to self-report segments in line with their management practice, but at the substantial cost of comparability over time and across firms.[10] Even in the pre-1997 period, Denis, Denis and Sarin (1997) document frequent arbitrary reporting changes in the number of segments which are unrelated to changes in business operations. Villalonga (2004b) finds that using data from the U.S. census to identify segments rather than data from Compustat results in a premium rather than a discount, a result potentially due to firms strategically manipulating their segment reporting to appear less valuable than a portfolio of single-segment competitors as is shown in Botosan and Stanford (2005).[11]

As a practical matter, inaccurate classification of industry codes to segments will contaminate the excess value measure, since by construction it is designed to rely exclusively on the accuracy of 4-digit SIC classification to match diversified firm segments and focused firms. A cursory comparison of the Segment and Industrial Compustat files' SIC code matches reveals that this is a potentially serious issue. Focused firms have different 4-digit SIC codes in the two files over 20% of the time; approximately 5% differ at even the 3-digit SIC code level. Over 34% of diversified firms have different SIC codes in the Industrial Annual file than their maximum sales segment in the Segment files.

[10] Among the many sources that document this effect are Berger and Hann (2003), Hund, Monk and Tice (2010), and Sanzhar (2006).

[11] Villalonga (2004b) and Montgomery (1994) also document the problems with the minimum segment reporting threshold of 10% of the total firm's sales. In particular, Montgomery (1994) examines the largest firms in the economy (as we do here) and finds that they are far more diversified than reported in the Compustat data. The fact that this additional diversification appears exactly in the set of firms we identify as having the highest premium strengthens the case for aggregate economic benefits to diversification.

In addition, market-to-sales ratios differ dramatically across 4-digit SIC codes. Table 5 depicts the averages and standard deviations of the market-to-sales ratio of the median focused firm from a 4-digit SIC code across years and 1-digit SIC codes. While the grand average market-to-sales ratio is 1.466, the standard deviation is 1.133, indicating that within 1-digit industries there are exceptionally large differences in market-to-sales ratios of median focused firms at the 4-digit SIC code level. Even median focused firms in 4-digit SIC codes adjacent to each other have very different market-to-sales ratios. Table 5 documents the absolute value of the difference in market-to-sales ratio across 4-digit SIC codes, again averaged over 1-digit SIC codes and years. Market-to-sales ratios of adjacent SIC codes are on average 0.724 different, nearly a 50% discrepancy from the mean level. The standard deviation of this difference is also large, indicating further instability in the relationship between market-to-sales values and 4-digit SIC codes.[12]

Taken together, these results cast serious doubts on the efficiency and reliability of constructing an excess-value measure matching segments solely on the dimension of the finest industry match. Even if the SIC code of either the diversified firm segment or the focused firm is misclassified by the minimum that it could be, the market-to-sales ratio of either could be biased by over 50%. Given these problems and the previously documented mismatch between focused firms and diversified firms on dimensions related to the uncertainty of their growth rates, we propose an alternative methodology for measuring the cost of diversification to the firm, and by extension, to the economy as a whole.

[12] Hoberg and Phillips (2011) find that their text-based measure of industry classification is better than conventional measures as an indicator of the true level of competition between firms and of related firm characteristics, such as profitability.

6. Matching Estimators of the Diversification Discount

If the diversification discount is driven by differences in the uncertainty of growth rates between focused and diversified firms, then controlling for that uncertainty should cause the differences in market value between them to disappear. Both Bevelander (2002) and Borghesi, Houston, and Naranjo (2007) find that matching on age within tight industry matches can reduce the diversification discount as measured by the excess value measure by 30% to 40%. Both note, however, the difficulty with maintaining a reasonable age match while preserving the tight industry match for the segments and adopt wide ranges for age to solve this problem.

We adopt a different approach to the problem. Rather than preserving the problematic industry matches for each segment, we compare focused and diversified firms directly, controlling for characteristics that are correlated with growth rate uncertainty as in Pastor and Veronesi (2003). The intuition is simple: if the diversified corporate form is systematically destroying firm value via agency costs, empire building, or inefficient cross-subsidization, then the difference between the market values of focused and diversified firms should be magnified by matching on age and sales. Large and old focused firms should be worth far more than their diversified counterparts. But if the measured discount is purely an artifact of comparing firms with different growth rate uncertainty, then the difference between market values of focused and diversified firms should disappear as we match on age and sales. By matching apples with apples, we should be able to straightforwardly identify the rotten. Using two completely different matching paradigms, one parametric and one non-parametric, we show in the following sections that the conventionally measured diversification discount is entirely an artifact of mismatching on elements directly correlated with growth rate uncertainty.

6.1 Propensity Score Matching Estimators

Propensity score matching provides a method for controlling confounding characteristics in an observational, rather than experimental, context. Specifically, to isolate the effect of organizational form on firm value we must control for variables such as sales, age, and volatility that are correlated both with firm value and with organizational form. These variables jointly proxy for the uncertainty of growth rates in the model of Pastor and Veronesi (2003), and they demonstrate that such uncertainty is correlated with market values. Hund, Monk, and Tice (2010) show that many empirical facts associated with diversification can be explained by interpreting the diversification discount as a difference in uncertainty of growth rates between diversified and focused firms. If diversified firms have lower growth rate uncertainty than focused firms, then it is critical to control for this difference in assessing the effects of organizational form on firm value.

Models using propensity score methods to control for endogeneity in the diversification decision have been used previously, most notably in Villalonga (2004a) and Colak (2010). Both of these papers model the decision to diversify directly on a restricted sample of firms that are moving from focused status to diversified status.[13] Our use of propensity scores is more limited, and admittedly more ad hoc; we are primarily interested in using the propensity scores as a parametric balancing metric to summarize multiple dimensions of potentially confounding covariates. Later in this section we supplement the results from this matching algorithm with results from a non-parametric balancing metric, coarsened exact matching, to test whether our results are driven by the particular form of matching estimators we choose.[14]

[13] Colak (2010) also considers the decision to spin-off, or re-focus as a function of endogenous firm characteristics.

[14] An excellent summary of matching estimators which discusses both propensity scores and coarsened exact matching can be found in Stuart (2010).

To estimate propensity scores, we regress diversification status on our proxies for growth rate uncertainty (sales, age, and return volatility) using a probit specification, and compute the propensity score as the predicted value of the regression.[15] We then match each diversified firm to the closest 3 neighboring focused firms based on their estimated propensity scores, and form the weighted average of their market value (equity market value plus book value of debt) to compare to the matched diversified firm.[16]

Table 6 presents the results of the propensity score matching procedure for four cases: treating each firm-year observation as independent; matching separately for each 1-digit industry; matching separately for each year; and matching separately within each year and 1-digit SIC code pair. Columns 1 and 2, where firms can be matched across years, are only accurate under unrealistic stationarity conditions and are primarily presented as benchmarks to compare the accuracy of the matching algorithm in a more realistic yearly context. Table 6 also highlights the importance of controlling for these characteristics since as emphasized in Section 2, diversified firms are much older, have dramatically higher sales, and lower return volatility than focused firms. Matching within year (and then again within both industry and year) substantially improves the average discrepancy between diversified and focused firms as compared to the unmatched sample, although sales are still significantly different (showing the difficulty in comparing the two organizational forms along common dimensions). One of our main results is the average effect on market value documented in Table 6. Once we control for variables correlated with growth rate uncertainty, diversified firms have an average market value of 458 million more than focused firms when we restrict years to match and an average market value of 280 million more than focused firms when we restrict matches to be in the same year

[15] Our results do not depend on the particular form of the propensity score regression, and hold whether we compute propensity scores using a logit, or as the odds ratio.

[16] Using five neighbors or a slightly different matching criterion does not alter our results.

and industry. Consistent with the aggregate economic benefits documented in Section 3, diversified firms add value at the firm level once the characteristics associated with growth rate uncertainty are balanced.

6.2 Coarsened Exact Matching Estimators

Common criticisms of propensity score methods are that they are prone to model mis-specification in the propensity score estimation and that they are focused on achieving optimal average balance among the covariates and not balance over their entire distributions (including ensuring correct common support). Coarsened Exact Matching (CEM) is a nonparametric technique developed in Iacus, King, and Porro (2011a) that ensures common support and bounds on the maximum imbalance between the covariate distributions of the diversified and focused firms. King et al. (2011) also show that CEM methods dominate most other matching methods in reducing estimation error, bias, and model dependence.

Essentially, CEM generates a multi-dimensional histogram of the covariates to match upon, dividing each variable into multiple bins (potentially of varying widths). The intersection of the "top" bin (for example) of all covariates forms a strata, or area associated with the highest values of all covariates. All possible intersections of the bins are computed and this forms the total set of strata for the matching procedure. Once formed, all diversified firm observations in a particular strata are matched with the focused firms in each strata and the weighted averages on the outcome variable (in this case, the market value of the firm) constitute the "treatment" effect. Any strata with no focused or no diversified firms are discarded (along with the firm observations in them) ensuring that the matching is only on the common support of the distributions.[17]

[17] Details on the statistical properties of the estimators, including the bounds on error and model dependence, are available in Iacus, King, and Porro (2011).

Implementation of CEM estimators involves a choice of "bin" size; too wide a bin size results in inefficient matching, whereas too narrow of a "bin" may result in discarding too many observations. For our data we use a simple rule based on the range of the data for age and volatility, and an optimization-based rule to select bins for the highly skewed and multi-modal sales distribution.[18]

Table 7 presents our results using CEM to match sales, age and return volatility and examine the effects of diversification status on firm value. The first column shows that the mismatch between focused and diversified firms on age, sales, and return volatility is persistent not just for the mean values, but all along the quartiles of each distribution. For example, the median diversified firm is 14 years older, has 195 million more in sales, and 1.95% lower return volatility than its match from focused firms. Also apparent is the extreme skewness of the sales distribution whose difference in mean is above the difference at the 75th percentile. As in Table 6, columns headed Unconstrained and Industry are provided primarily as a benchmark, since they allow matching across years which requires unrealistically strong assumptions about firm-year independence and stationarity. Columns headed Year and Industry-Year, however, present strong evidence that once correctly balanced on proxies for growth rate uncertainty, there is no diversification discount. Balanced fairly well across the mean and all the quartiles of the covariates and requiring matches to be in the same year, diversified and focused firm values are virtually identical. When we constrain the matches to match exactly on year and 1-digit SIC code, and match across the entire distributions of sales, age, and return volatility, diversified firms are more valuable than focused firms.

[18] Specifically, we use Sturges' rule for the smoothly distributed age and volatility distributions and the methods developed for multi-modal and skewed data in Shimazaki and Shinomoto (2007) for the sales variable distribution.

Using both model dependent (propensity score) and more non-parametric (CEM) matching estimators, we find that diversified firms either have a similar value or are more valuable than focused firms once we control for sales, age, and volatility (and require matches to be in the same year and wide industry group). Matching apples to apples, diversification does not destroy value.

7. Conclusion

Since the seminal paper of Berger and Ofek (1995), a host of theories have evolved to explain why diversified firms are worth less than an imputed firm value derived from median industry matched focused firm market-to-sales ratios. In this paper we show that this difference is primarily driven by two things: matching older, larger, and less volatile firms with younger, smaller, and more volatile firms; and a large number of economically small firms which trade at a discount with this particular metric. Diversified firms that trade at a premium also have the largest market value, which changes the equal-weighted discount often reported in the literature to a value-weighted premium. In aggregate economic terms, diversification creates value. We also argue that controlling for industry in the measure of excess value to the exclusion of size, age, and volatility amounts to comparing apples to oranges. Hund, Monk, and Tice (2010) show that many empirical facts about diversification (a discount in levels, positive changes in values for diversified firms, higher idiosyncratic volatility for focused firms, and discounts which co-vary with the business cycle) can be explained by interpreting the diversification discount as matching firms with low uncertainty about growth rates (diversified firms) with firms with high uncertainty about growth rates (focused firms). In this paper, we show that once the uncertainty environment is correctly controlled using either parametric or non-parametric matching methods, there is no diversification discount. Indeed, it is difficult to reconcile our finding that the oldest, largest, and most valuable firms in the economy are also diversified with agency cost theories of

the firm that suggest such firms will be most prone to the vagaries of empire-building, inefficient investment, and rent-seeking by managers. It may be fruitful to now explore the potential dimensions and contexts along which the conglomerate form adds value, such as advantages in labor markets (as in Tate and Yang (2011)), alleviating credit constraints (as in Dimitrov and Tice (2006)), compensation for better performance during economic crises (as in Kuppuswamy and Villalonga (2010)), or higher product differentiation (as in Hoberg and Phillips (2012)). We leave this for further research.

References

Berger, Philip and Rebecca Hann, 2003, The impact of SFAS no. 131 on information and monitoring, *Journal of Accounting Research* 41, 163-223.

Berger, Philip and Eli Ofek, 1995, Diversification's effect on firm value, *Journal of Financial Economics* 37, 39-65.

Bevelander, Jeffrey, 2002, Tobin's q, corporate diversification, and firm age, Unpublished Dissertation, M.I.T.

Borghesi, Richard, Joel Houston, and Andy Naranjo, 2007, Value, survival, and the evolution of firm organizational structure, *Financial Management* 36, 5-31.

Botosan, Christine A. and Mary Stanford, 2005, Managers' motives to withhold segment disclosures and the effect of SFAS No. 131 on Analysts' Information Environment, *The Accounting Review* 80, 751-771.

Campa, Jose and Simi Kedia, 2002, Explaining the diversification discount, *Journal of Finance* 57, 1731-1762.

Colak, Gonul, 2010, Diversification, refocusing, and firm value, *European Financial Management* 16, 422-448.

Denis, David, Diane Denis, and Atulya Sarin, 1997, Agency problems, equity ownership, and corporate diversification, *Journal of Finance* 52, 135-160.

Dimitrov, Valentin and Sheri Tice, 2006, Corporate diversification and credit constraints: real effects across the business cycle, *Review of Financial Studies* 19, 1465-1498.

Fink, Jason, Kristin Fink, Gustavo Grullon, and James Weston, 2012, What drove the increase in idiosyncratic volatility during the internet boom?, *Journal of Financial and Quantitative Finance*, forthcoming.

Graham, John, Michael Lemmon, and Jack Wolf, 2002, Does corporate diversification destroy value?, *Journal of Finance* 57, 695-720.

Hoberg, Gerald and Gordon Phillips, 2011, Text-based network industries and endogenous product differentiation, Working Paper.

Hoberg, Gerald and Gordon Phillips, 2012, Conglomerate industry choice and product differentiation, Working Paper.

Hoechle, Daniel, Markus Schmid, Ingo Walter, and David Yermack, 2011, How much of the diversification discount can be explained by poor corporate governance?, Working Paper.

Hund, John, Donald Monk, and Sheri Tice, 2010, Uncertainty about average profitability and the diversification discount, *Journal of Financial Economics* 96, 463-484.

Iacus, Stefano, Gary King, and Giuseppe Porro, 2011a, Multivariate matching methods that are monotonic imbalance bounded, *Journal of the American Statistical Association* 106, 345-361.

Iacus, Stefano, Gary King, and Giuseppe Porro, 2011b, Causal inference without balance checking: coarsened exact matching, *Political Analysis*, forthcoming.

Jovanovic, Boyan and Peter Rousseau, 2001, Why wait? A century of life before IPO, *American Economic Review* 91, 336-41.

Khanna, Naveen and Sheri Tice, 2001, The bright side of diversification, *Journal of Finance* 56, 1489-1528.

King, Gary, Richard Nielsen, Carter Coberley, James Pope, and Aaron Wells, 2011, Comparative effectiveness of matching methods for causal inference, Working Paper.

Kuppuswamy, Venkat and Belen Villalonga, 2010, Does diversification create value in the presence of external financing constraints? Evidence from the 2007-2009 financial crisis, Working Paper.

Lamont, Owen and Christopher Polk, 2002, Does diversification destroy value? Evidence from industry shocks, *Journal of Financial Economics* 63, 51-77.

Maksimovic, Vojislav and Gordon Phillips, 2002, Do conglomerate firms allocate resources inefficiently across industries? Theory and evidence, *Journal of Finance* 57, 2167-2183.

Montgomery, Cynthia, 1994, Corporate diversification, *Journal of Economic Perspectives* 8, 163-178.

Ozbas, Oguzhan and David Scharfstein, 2010, Evidence on the dark side of internal capital markets, *Review of Financial Studies* 23, 581-599.

Pastor, Lubos and Pietro Veronesi, 2003, Stock valuation and learning about profitability, *Journal of Finance* 58, 1749-1789.

Rajan, Raghuram, Henri Servaes, and Luigi Zingales, 2000, The cost of diversity: the diversification discount and inefficient investment, *Journal of Finance* 46, 3-27.

Sanzhar, Sergey, 2006, Discounted but not diversified: organizational structure and conglomerate discount, Working Paper.

Scharfstein, David and Jeremy Stein, 2000, The dark side of internal capital markets: divisional rent-seeking and inefficient investment, *Journal of Finance* 55, 2537-2564.

Shimazaki, Hideaki and Shigeru Shinomoto, 2007, A method for selecting the bin size of a time histogram, *Neural Computation* 19, 1503-27.

Shin, Hyun-Han. and Rene Stultz, 1998, Are internal capital markets efficient?, *Quarterly Journal of Economics* 113, 531-552.

Stuart, Elizabeth, 2010, Matching methods for causal inference: a review and a look forward, *Statistical Science* 25, 1-21.

Tate, Geoffrey and Liu Yang, 2011, The bright side of diversification: evidence from internal labor markets, Working Paper.

Villalonga, Belen, 2004a, Does diversification cause the "diversification discount"?, *Financial Management* 33, 5-27.

Villalonga, Belen, 2004b, Diversification discount or premium? New evidence from the business information tracking series, *Journal of Finance* 59, 479-506.

Table 1

Summary Statistics by Diversification Status and Excess Value

This table presents summary statistics for focused and diversified firms by premium and discount status from 1977-2009. Using the Compustat segment file, firms are categorized into focused (one business segment) and diversified (greater than one business segment). Firms with an excess value measure, defined in Berger & Ofek (1995) as the log ratio of total firm capitalization (*V*) to imputed firm capitalization, that is greater than or equal to zero are categorized as premium firms while those with excess value less than zero are discount firms. *V* is the total capitalization of the firm calculated as *MVE* plus the book value of debt. *MVE* is the firm fiscal year-end market value of equity from CRSP. Assets, Sales, and Profit Margin (ebitda/sales) are from Compustat data, Age is firm age calculated using data and methods from Jovanovich & Rousseau (2002), and Return Volatility and Excess Return Volatility are calculated respectively as the standard deviation of monthly returns and monthly returns net of the market index from CRSP data. Median values are in italics below mean values. All numbers are in millions except for Age and those shown as percentages.

	Focused			**Diversified**		
	Discount	**Premium**	**Total**	**Discount**	**Premium**	**Total**
MVE	469	1,522	1,012	931	3,664	2,072
	60	*236*	*121*	*105*	*422*	*185*
V	715	1,851	1,301	1,489	4,388	2,699
	93	*300*	*173*	*177*	*592*	*293*
Assets	861	1,162	1,016	2,080	2,724	2,349
	107	*191*	*144*	*219*	*459*	*299*
Sales	971	1,079	1,026	1,998	2,559	2,232
	142	*179*	*160*	*304*	*459*	*354*
Age	27	26	26	43	48	45
	17	*16*	*17*	*30*	*33*	*31*
Earnings	52	93	73	90	242	154
	3	*10*	*6*	*8*	*27*	*13*
Profit Margin	8.23%	13.50%	10.90%	9.70%	14.30%	11.60%
	7.70%	*13.60%*	*10.40%*	*9.15%*	*13.90%*	*10.80%*
Return Volatility	14.80%	13.90%	14.30%	12.50%	12.00%	12.30%
	12.90%	*12.10%*	*12.50%*	*10.80%*	*10.20%*	*10.50%*
Excess Return Volatility	14.10%	13.00%	13.50%	11.50%	10.90%	11.30%
	12.20%	*11.10%*	*11.60%*	*9.71%*	*9.03%*	*9.43%*
N	26,599	28,364	54,963	14,085	10,094	24,179

Table 2

Mean and Median Excess Values by Decile

This table shows mean and median excess values (EV) computed as in Berger & Ofek (1995) for deciles of MVE (Panel A), Sales (Panel B), and Age (Panel C) for focused and diversified firms over the years 1977-2009. Using the Compustat segment file, firms are categorized into focused (one business segment) and diversified (greater than one business segment). Decile ranks are computed using the previous year's values for the variables, and formed within focused and diversified groups. *MVE* is the firm fiscal year end market value of equity from CRSP, Sales are taken directly from Compustat, and Age is firm age calculated using methods and data from Jovanovich & Rousseau (2002).

Panel A: MVE

Focused

Decile	N	Mean MVE	Mean EV	Median EV
1	4477	18	-0.414	-0.419
2	4460	42	-0.294	-0.278
3	4465	71	-0.197	-0.173
4	4459	107	-0.109	-0.074
5	4458	160	-0.031	0.000
6	4465	244	0.030	0.000
7	4462	377	0.090	0.055
8	4462	624	0.149	0.118
9	4463	1,242	0.200	0.181
10	4447	8,159	0.284	0.266
Total	44618	1,102	-0.029	0.000

Diversified

Decile	N	Mean MVE	Mean EV	Median EV
1	2026	20	-0.459	-0.495
2	2015	49	-0.312	-0.342
3	2018	90	-0.213	-0.249
4	2013	156	-0.157	-0.189
5	2013	243	-0.137	-0.155
6	2019	401	-0.043	-0.066
7	2017	675	-0.014	-0.019
8	2014	1,278	0.023	0.012
9	2019	2,716	0.070	0.058
10	1996	16,674	0.182	0.191
Total	20150	2,215	-0.107	-0.120

Panel B: Sales

Focused

Decile	N	Mean Sales	Mean EV	Median EV
1	4477	33	-0.063	-0.030
2	4460	51	-0.076	-0.047
3	4466	79	-0.089	-0.066
4	4458	118	-0.079	-0.049
5	4458	173	-0.061	-0.017
6	4465	257	-0.001	0.000
7	4462	396	-0.018	0.000
8	4462	653	0.012	0.000
9	4463	1,269	0.016	0.000
10	4447	8,445	0.066	0.040
Total	44618	1,145	-0.029	0.000

Diversified

Decile	N	Mean Sales	Mean EV	Median EV
1	2026	38	-0.177	-0.225
2	2015	79	-0.172	-0.196
3	2018	138	-0.144	-0.161
4	2013	230	-0.138	-0.173
5	2013	363	-0.135	-0.173
6	2019	579	-0.108	-0.124
7	2017	978	-0.058	-0.064
8	2014	1,805	-0.052	-0.052
9	2019	3,731	-0.026	-0.026
10	1996	15,489	-0.055	-0.030
Total	20150	2,329	-0.107	-0.120

Panel C: Age

Focused

Decile	N	Mean Age	Mean EV	Median EV
1	3875	3	-0.022	0.000
2	3791	6	-0.026	0.000
3	4463	9	-0.001	0.000
4	4196	12	-0.012	0.000
5	4469	16	-0.032	-0.002
6	4667	20	-0.025	0.000
7	4561	25	-0.042	-0.007
8	4753	34	-0.086	-0.057
9	4901	52	-0.040	-0.004
10	4901	94	-0.004	0.000
Total	44577	29	-0.030	0.000

Diversified

Decile	N	Mean Age	Mean EV	Median EV
1	1814	4	-0.186	-0.201
2	1824	10	-0.117	-0.147
3	2156	15	-0.126	-0.150
4	1770	21	-0.135	-0.154
5	2044	28	-0.117	-0.158
6	2011	38	-0.105	-0.102
7	2122	52	-0.106	-0.118
8	2118	69	-0.108	-0.145
9	2152	91	-0.083	-0.087
10	2127	131	0.001	0.011
Total	20138	48	-0.106	-0.120

Table 3

Equal- and Value-Weighted Excess Value Averages by Year

This table presents results for the excess value (EV) as computed in Berger & Ofek (1995) for diversified firms averaged by year, where averages are the simple average over all diversified firms (Eq. wtd) and value-weighted over diversified firms by MVE and *V*. MVE is the firm fiscal year end market value of equity from CRSP. *V* is the total capitalization of the firm calculated as MVE plus the book value of debt. Loss/Gain Eq. wtd (*V* wtd) is computed as the equal (*V*) weighted premium or discount multiplied by the value of the firm, *V*. The percentage of diversified firms with an excess value measure greater than zero in a particular year is provided in the % Prem column. The last row of the table provides summary results after including financial firms (SIC codes 6000-6999) in the sample.

		Mean (EV)				Loss/Gain		
Year	N	Eq. wtd	MVE wtd	*V* wtd	*V*	Eq. wtd	*V* wtd	% Prem
1977	841	-0.0690	0.3223	0.2218	436	-30	141	42%
1978	975	-0.0835	0.1956	0.1255	423	-35	83	42%
1979	962	-0.1233	0.0919	0.0249	489	-60	45	38%
1980	907	-0.1223	0.1133	0.0396	603	-74	68	39%
1981	891	-0.1418	0.1536	0.0701	592	-84	91	39%
1982	856	-0.1717	0.0981	0.0349	642	-110	63	37%
1983	822	-0.1778	0.0120	-0.0238	800	-142	10	35%
1984	801	-0.1460	0.0322	-0.0054	800	-117	26	38%
1985	786	-0.1340	0.1004	0.0523	998	-134	100	37%
1986	713	-0.0943	0.1698	0.1224	1,261	-119	214	41%
1987	655	-0.0781	0.2459	0.1668	1,350	-105	332	44%
1988	638	-0.1015	0.1772	0.0832	1,349	-137	239	40%
1989	588	-0.0923	0.3654	0.2552	1,642	-152	600	44%
1990	565	-0.1019	0.4183	0.2867	1,712	-174	716	43%
1991	566	-0.1266	0.3995	0.2494	2,175	-275	869	40%
1992	588	-0.0916	0.3212	0.1915	2,129	-195	684	44%
1993	582	-0.1061	0.1964	0.0740	2,209	-234	434	43%
1994	606	-0.1361	0.1878	0.0721	1,962	-267	368	40%
1995	604	-0.1224	0.2784	0.2089	2,219	-272	618	42%
1996	589	-0.0816	0.2957	0.2305	2,804	-229	829	46%
1997	561	-0.0922	0.4043	0.3523	3,566	-329	1442	45%
1998	901	-0.0124	0.7401	0.5873	3,560	-44	2635	48%
1999	1,000	-0.0584	0.5279	0.4601	4,240	-247	2238	47%
2000	919	-0.0902	0.4471	0.3513	3,975	-359	1777	44%
2001	832	-0.1186	0.3442	0.2436	4,435	-526	1526	42%
2002	799	-0.0642	0.3984	0.2862	3,815	-245	1520	47%
2003	749	-0.1502	0.3206	0.2382	5,276	-792	1691	41%
2004	733	-0.1458	0.3221	0.2123	5,748	-838	1851	41%
2005	711	-0.1284	0.2482	0.1845	5,466	-702	1357	42%
2006	692	-0.1412	0.2034	0.1162	6,637	-937	1350	39%
2007	630	-0.1022	0.2484	0.1850	7,249	-741	1801	42%
2008	635	-0.0515	0.4193	0.3304	5,391	-277	2260	47%
2009	513	-0.1219	0.1866	0.1481	6,512	-794	1215	43%
Total	24,210	-0.1082	0.3272	0.2395	2,802	-296	885	42%
w/Fin.	28,501	-0.1081	0.2937	0.2168	3,497	-375	975	42%

Table 4

S&P 500 Membership and Percentage of S&P 500 Market Value by Diversification Status

This table presents the number and percentage of total market value for diversified and focused firms in the S&P 500 index by year. Using the Compustat segment file, firms are categorized into focused (one business segment) and diversified (greater than one business segment). Total S&P 500 firms are less than 500 due to absence from the Compustat segment file. Firm market values of equity (MVE) are computed using their average monthly market value of equity over the year from the CRSP monthly file, and the columns labeled "% of MVE" are percentages of total S&P500 market value computed from all firms with data for that year.

		Focused		Diversified	
Year	S&P Firms	N	% of MVE	N	% of MVE
1977	462	106	23.2%	356	76.8%
1978	465	109	23.3%	356	76.7%
1979	467	111	21.3%	356	78.7%
1980	466	107	17.1%	359	82.9%
1981	466	112	19.0%	354	81.0%
1982	467	114	26.2%	353	73.8%
1983	460	116	29.2%	344	70.8%
1984	470	131	27.9%	339	72.1%
1985	462	135	32.9%	327	67.1%
1986	460	142	31.6%	318	68.4%
1987	462	152	31.2%	310	68.8%
1988	454	151	27.7%	303	72.3%
1989	458	151	28.2%	307	71.8%
1990	460	156	27.0%	304	73.0%
1991	461	156	27.8%	305	72.2%
1992	459	164	29.1%	295	70.9%
1993	461	169	29.7%	292	70.3%
1994	460	180	33.9%	280	66.1%
1995	459	188	37.1%	271	62.9%
1996	462	203	41.1%	259	58.9%
1997	456	206	43.4%	250	56.6%
1998	456	117	23.2%	339	76.8%
1999	461	100	17.9%	361	82.1%
2000	461	113	21.2%	348	78.8%
2001	463	112	17.8%	351	82.2%
2002	458	106	16.4%	352	83.6%
2003	459	113	19.6%	346	80.4%
2004	463	114	18.4%	349	81.6%
2005	462	107	17.6%	355	82.4%
2006	465	101	15.5%	364	84.5%
2007	467	106	15.6%	361	84.4%
2008	465	106	16.4%	359	83.6%
2009	437	100	20.2%	337	79.8%
Average	461	132	25.1%	329	74.9%

Table 5

Differences in Market-to-Sales in Adjacent SIC Codes

This table reports the summary statistics of the market-to-sales ratio of the median focused firm in every 4-digit SIC code over the period 1977-2009 contained in our matched CRSP/Compustat/Segment data. Market-to-sales is calculated as the ratio of total capital (market value of end-of-fiscal year equity plus book value of debt) to total sales. Columns 1 and 2 report the mean and standard deviation of the market-to-sales ratio for the median focused firm averaged over 1-digit SIC codes and years as a benchmark. Columns 3 and 4 report the mean and standard deviation of the absolute value of the difference in median market-to-sales ratios across adjacent (ascending) 4-digit SIC codes, averaged across 1-digit SIC codes and all years.

	Market-to-Sales		Adjacent Differences (4 digit)	
SIC 1-Digit	**Mean**	**Std. Dev**	**Mean**	**Std. Dev**
0	2.228	1.356	1.375	1.119
1	1.878	1.775	1.131	1.273
2	1.064	0.962	0.682	0.918
3	1.052	0.663	0.585	0.589
4	2.147	1.766	1.472	1.621
5	0.542	0.334	0.350	0.427
7	1.725	1.133	1.227	1.142
8	1.646	1.234	1.207	1.378
All	1.466	1.133	0.724	0.811

Table 6

Diversification Effects on Value Using Propensity Matched Firms

This table reports mean values for various matching characteristics (Sales, Age, and Return Volatility) for diversified firms and focused firm matches along with the p-value of the test of the statistical significance of the differences between the mean values. Sample average treatment effects on value, *V*, are presented in the row titled "diff" along with the standard error and p-value of the effect. Diversified firms are matched with focused firms using propensity scores from probit regressions. Diversified firms are compared against their 3 closest matched neighbors with respect to the generated propensity score. Separate columns contain results from matching diversified firms on sales, age, and annual return volatility using no additional constraints, constraining matches to be within the same one-digit SIC code industry, the same year, and the same industry and year. Sales are computed from Compustat, Age is calculated using methods and data from Jovanovich & Rousseau (2002), and Return Volatility is the standard deviation of monthly returns. *V* is the total capitalization of the firm calculated as the firm fiscal year-end market value of equity from CRSP plus the book value of debt. Data span the years 1977-2009.

	Unmatched	Unconstrained	Industry	Year	Industry-Year
Sales					
Diversified	2,231	2,232	2,232	2,232	2,232
Focused	1,036	2,221	1,945	1,994	1,821
p-value		0.8947	0.0001	0.0013	0.0000
Age					
Diversified	45	45.1	45.1	45.1	45.1
Focused	26	45.1	45.8	45.4	45.2
p-value		0.9174	0.0534	0.3720	0.8061
Volatility					
Diversified	12.28%	12.28%	12.28%	12.28%	12.28%
Focused	14.35%	12.29%	12.39%	12.32%	12.37%
p-value		0.8917	0.1254	0.5859	0.1844
V					
Diversified	2,697	2,699	2,699	2,699	2,700
Focused	1,313	2,427	2,538	2,241	2,420
diff.		272	161	**458**	**280**
s.e.		*(99.9)*	*(102.4)*	***(96.7)***	***(102.5)***
p-value		2.7222	0.1164	**0.0000**	**0.0062**

Table 7

Diversification Effects on Value Using Coarsened Exact Matching

This table reports sample average treatment effects on value, *V*, for diversified firms matched with focused firms using the coarsened exact matching algorithm from Iacus, King, and Porro (2011) for data from 1977-2009. Diversified firms are compared against all focused firms within their coarsened strata, with weights computed proportionally within the strata and strata defined using the bin selection algorithm in Shimazaki and Shinomoto (2007). Table entries for sales, age, and return volatility reflect differences between diversified and focused firms at the means and the 25th (p25), 50th (p50), and 75th (p75) percentiles. Separate columns contain results from matching diversified firms on sales, age, and annual return volatility using no additional constraints, constraining matches to be within the same one-digit SIC code industry, the same year, and the same industry and year. Sales are computed from Compustat, Age is calculated using methods and data from Jovanovich & Rousseau (2002), and Return Volatility is the standard deviation of monthly returns. *V* is the total capitalization of the firm calculated as the firm fiscal year-end market value of equity from CRSP plus the book value of debt.

	Unmatched	Unconstrained	Industry	Year	Industry-Year
Sales					
mean	1195	156	187	188	204
p25	42	20	23	25	25
p50	195	107	113	110	108
p75	995	530	530	496	435
Age					
mean	18.63	0.26	0.26	0.24	0.25
p25	5.00	0.00	0.00	0.00	0.00
p50	14.00	0.00	1.00	0.00	1.00
p75	36.00	0.00	0.00	0.00	0.00
Return Volatility					
mean	-2.07%	-0.22%	-0.25%	-0.24%	-0.22%
p25	-1.37%	-0.26%	-0.30%	-0.28%	-0.26%
p50	-1.95%	-0.33%	-0.35%	-0.30%	-0.28%
p75	-2.65%	-0.23%	-0.24%	-0.21%	-0.15%
V					
mean	1384.5	53.1	-183.2	**61.9**	**89.7**
s.e	*(64.6)*	*(64.0)*	*(54.5)*	***(43.7)***	***(26.9)***
p-value	0.0000	0.4060	0.0010	**0.1560**	**0.0010**