Systematic Heterogeneity versus Average Effects in the Returns to Diversification

Bharat N. Anand* Dmitri Byzalov†

October 31, 2011

Abstract

Differences across firms are at the root of theories of strategic advantage. In this paper we examine empirically systematic differences across firms in the context of a particular debate – the returns to diversification. We focus on isolating the systematic part of variation in returns to diversification across firms, and how large it is relative to the average diversification discount or premium (the focus of prior literature). Our empirical approach carefully decomposes overall variation in performance into possibly-persistent but ultimately transitory shocks, systematic heterogeneity in the returns to diversification, and generic heterogeneity from all other sources. Using data on excess values for Compustat firms over three decades, we find that the systematic heterogeneity in returns to diversification is large, both in absolute magnitude and relative to the mean discount. For example, the standard deviation of this systematic heterogeneity amounts to 18-34% of market value of the firm, and is roughly twice as large as the mean discount; and, about one-third of all firms actually display a systematic "diversification premium". Our estimates imply that the mean discount accounts for only 9-19% of total systematic variation in excess values due to diversification. The results are robust to the inclusion of standard observables employed in prior work, and

*Harvard Business School, Soldiers Field Road, Boston, MA 02163. Phone: (617) 495-5082; Fax: (617) 495-0355; email: banand@hbs.edu.

†Fox School of Business, Temple University, Philadelphia, PA 19122. Phone: (215) 204-3927; Fax: (215) 204-5587; email: dbyzalov@temple.edu.

to the endogeneity of diversification choices. These results have important theoretical and prescriptive implications for research on diversification, and suggest that shifting the focus "from means to variances" is likely to be fruitful in other areas of empirical strategy research.

1. Introduction

Differences across firms are at the root of competitive advantage. Furthermore, these differences can persist for many reasons, as a large body of theory now illustrates. For example, firms may acquire and possess hard-to-imitate resources (Wernerfelt 1984, Barney 1991, Peteraf 1993). Or, firm choices may be hard to mimic because of complementarities across them that result in system complexity (for eg., Porter 1996, Rivkin 2000) or causal ambiguity (Lippman and Rumelt 1982, Ryall 2009). One important consequence of these theories is that the costs and benefits of the same decision may greatly vary across firms, and these differences may persist year after year.

Despite rich progress in the theories of strategic advantage over the last twenty years, empirical research has by and large not taken differences as seriously as the theory would imply. Consider, for example, the performance consequences of any major firm level decision such as the decision to ally, the decision to diversify, the decision to locate abroad, or the decision regarding R&D investment. Vast literatures on each of these topics exist: for example, Lang and Stulz (1994), Berger and Ofek (1995), Campa and Kedia (2002) and Villalonga (2004a, 2004b) on the returns to diversification; Anand and Khanna (2000), Kale, Dyer, and Singh (2001) and Sampson (2004, 2007) on the value from alliances; Chan et al (1995) on the value from firm relocations; Prezas et al (2010) on the returns to off-shoring; Brickley and van Drunen (1990) on the effects of restructuring; Cockburn and Griliches (1988), Bernstein and Nadiri (1988), Erickson and Jacobson (1992), Hall (1993), Chan et al (2001) and Cassiman and Veugelers (2006) on the returns to R&D. The central focus of virtually all these papers, however, is the average "performance" effect of the particular policy variable (after conditioning on observables) rather than its differential impact across firms.[1] This is a surprising omission for both positive and normative reasons. First, if one were to take seriously the theory of interdependencies across choices, then the marginal effects of such decisions should vary across firms. Second, to the extent that heterogeneous effects are important in the data after controlling for observable differences, prescriptions drawn from average effects—for example, concerning whether or not diversification is value creating, whether firms should ally or not, or whether offshoring is beneficial—are less meaningful.

This paper attempts to shed light empirically on the relative importance of "means versus variances" in the context of a particular debate—the returns to diversification. The effect of diversification on firm value is perhaps the most studied question about multi-business firms, and the "diversification discount" the most well-known stylized fact to emerge from this literature. However, as several authors note, prescriptive implications of findings regarding the discount remain unclear because "all statements about the diversification discount are statements about averages. But in many ways, it is the variation that matters most to (those) who care about the normative implications of our research" (Gertner 2004). Some papers, starting with Lang and Stulz (1994), provide evidence regarding cross-sectional heterogeneity in the diversification discount. But heterogeneity in excess values, by itself, need not alter the prescriptive consequences of findings concerning mean differences for at least three reasons. First, cross-sectional heterogeneity may reflect idiosyncratic shocks rather than genuine systematic differences across firms. Second, even persistent differences in excess values may reflect temporary shocks that only slowly erode, rather than due to systematic heterogeneity. Third, even if systematic heterogeneity exists, its empirical magnitude may be small compared to differences in mean excess values.

A central question in this context, then, is: how large is the systematic heterogeneity in returns to diversification—both in absolute magnitude, and relative to the mean diversification discount as well as relative to other sources of systematic heterogeneity?

[1] Some studies allow for differential effects based on interaction with other observable variables. For example, in the context of returns to R&D, Cockburn and Griliches (1988) include interactions with the degree of patent protection, Bernstein and Nadiri (1988) interactions with industry dummies, and Cassiman and Veugelers (2006) interactions with external knowledge acquisition.

Surprisingly little is known about the answer to this question—just as for the other contexts above. There are several reasons this question is important and interesting. If the systematic variation in returns to diversification were small compared with the diversification discount, then the latter remains, even for any individual firm, the most relevant predictor of the value consequences from diversification. However, if systematic variation were large, then evidence of an "average" discount will have less prescriptive relevance for a large group of firms - namely, those diversified firms that are systematically successful. Instead, one would want to understand what are the underlying drivers of this success, i.e., whether it is due to their unusually high systematic returns to diversification or due to their systematically superior performance on dimensions unrelated to diversification. More than that, it would suggest that even for poorly performing diversified firms, understanding the sources of their inferior performance is likely to be a more pragmatic and fruitful approach than prescriptions based on the average diversification discount or premium.[2]

Some recent work on the diversification discount attempts to go beyond documenting the average effect by studying how certain cross-sectional observables explain heterogeneity in the discount. However, from a theoretical standpoint, it would be surprising if observables captured all, or even meaningfully relevant, systematic variation in firm strategies. The reason is that if all variation were due to simple and easily measured observables (for eg., "having a presence in China" or "related-SIC expansion"), then, according to virtually any theory of strategic advantage, imitation should be straightforward and equilibrium variation unlikely. Aside from this, however, variation due to observables does not by itself imply that systematic heterogeneity even exists or is important, for several reasons. First, there may be variation in observables for a firm over time, so that the same firm can display a premium in one period and a discount the next. Second, the relative importance of the mean discount versus (explained) heterogeneity may still be very large. For example, if explained variance is 5% of the mean discount, then the diversification

[2] For example, their inferior performance may be due to unusually low systematic returns to diversification, in which case refocusing is likely to create value. Alternatively, it could be due to very poor performance along other dimensions unrelated to diversification, combined with above-average returns to diversification, in which case refocusing may further destroy value.

discount still remains the most important driver for prescriptions about whether or not to break up. Last, endogeneity of observables (for eg., unrelated diversification and negative excess values may both result from poor management) makes prescriptions unclear. For all these reasons, our primary focus here is on characterizing the total magnitude of systematic heterogeneity in returns to diversification, and total magnitude of systematic heterogeneity from other sources, without attempting to attribute them to specific observable or unobservable factors.[3]

Several recent studies question the findings on the mean diversification discount. Some question the causal relation implied in studies on the diversification discount. For example, firms might diversify only if they have become relatively unproductive in their core business (e.g., Gomes and Livdan 2004) – in which case firms already suffer from a discount before they diversify, not because of it. Campa and Kedia (2002) and Villalonga (2004a) offer evidence on this (following their findings, we explicitly control for endogeneity of diversification in robustness checks).[4] Others question the sanctity of the data used in the first place; notably, Villalonga (2004b) notes that many previous studies that document the diversification discount rely on data that contain measurement problems, which may distort the estimates of returns to diversification. Like the previous literature on the discount, these studies focus on the differences in means between the group of firms that diversify and those that do not. Our primary focus, in contrast, is on systematic heterogeneity in firm-specific returns from diversification, and how large they are compared to the between-group differences in means. Indeed, the earlier concerns about drawing prescriptive implications from 'statements about averages' apply regardless of whether there is an average "discount" or "premium." Furthermore, they apply equally even if the average causal effect of diversification is zero, i.e., diversification neither creates not destroys value for an *average* firm.[5]

[3]In section 5, however, we examine the importance of standard observables typically employed in the literature in explaining the estimated systematic heterogeneity.

[4]In a related paper, Chevalier (2004) shows that investment patterns that look like cross-subsidization between divisions (and that could explain lower values for diversified firms) were already present in the pairs of merging firms *before* the diversifying merger. In addition, she shows that stock market valuations for diversifying mergers are the same as for other mergers.

[5]In this case, the conclusion based on the mean discount (the focus of prior literature) would be that

The paper makes three contributions. First, we estimate the relative magnitude of systematic heterogeneity in returns to diversification vis-a-vis the mean discount, in order to assess the importance of systematic heterogeneity in returns. Second, empirically identifying the heterogeneous effects of diversification across firms is not straightforward since there are various sources of performance heterogeneity that may have little to do with diversification. For example, heterogeneity may be generalized and unobserved to the researcher (Apple's superior performance may be attributable to a general innovation capability rather than its specific diversification from computers to MP3 players, mobile phones, then tablets); or, it may be transitory but somewhat persistent in the data (for eg., Apple's advantage in MP3 players erodes over time, but only slowly). As a result, a second contribution here is to employ an estimation approach that separates these various effects—the heterogeneous impact across firms from a particular choice (diversification), unobserved systematic heterogeneity due to general capabilities or superior strategy, and "somewhat persistent" but ultimately transitory effects—and that can be generalized beyond the context of diversification. Last, we illustrate the consequences of mis-specification by identifying a list of "successful diversifiers" that emerge from our model estimates and show how the identities of firms on this list vary (somewhat dramatically) across the different approaches.

A number of studies examine the sources of performance differences across firms by decomposing the overall variation in performance into industry, corporate and business-unit effects (e.g., Schmalensee 1985, Rumelt 1991, McGahan and Porter 1997, 1999, 2002, 2003, 2005, Chang and Singh 2000, Bowman and Helfat 2001). Most of these studies capture the various sources of performance differences using fixed effects. Our paper differs from these studies in two important ways. First, in contrast to these studies, our focus is on the returns to a particular strategic choice (diversification), and especially the systematic heterogeneity in the effects of this choice. As we illustrate in section 2, the implications of heterogeneity in firm performance depend crucially not only on the total magnitude of

diversification does not matter. However, this conclusion would not be accurate if the systematic heterogeneity in firm-specific returns to diversification is large, since diversification would matter a lot for a large group of firms – those in both tails of the distribution of systematic returns.

the systematic part of this heterogeneity, but also on how this systematic part decomposes into systematic heterogeneity attributed to diversification vis-a-vis other factors. Second, the fixed-effects approach employed in typical studies is likely to overestimate the actual *systematic* part of overall heterogeneity from each source, since the estimates of the fixed effects capture not only the genuine systematic component of firm performance but also the average of transitory shocks over a relatively small number of observations per firm. Consequently, findings of a large significant variance of the fixed effects in these studies do not necessarily indicate systematic heterogeneity—instead, they could be driven entirely by idiosyncratic or somewhat persistent but ultimately transitory shocks.[6] As we illustrate in section 3, a fixed-effects approach would in fact dramatically overestimate the actual magnitude of systematic heterogeneity in our data.

We employ two data sources. First, we examine Compustat-based firm-level and segment-level data for 1978-1996. This is the longest time period for which segment-level data are available and reported in a uniform format (SFAS 14), and has been employed in most prior studies of diversification. Second, we construct a new dataset for 1998-2008, the period following a change in segment reporting requirements (adoption of SFAS 131). Since these two datasets are not directly comparable, we conduct the analysis separately. However, the more recent dataset is of particular interest to check robustness of our results on the earlier sample.

Our results show that there is substantial and systematic heterogeneity across firms in the value from diversification - the standard deviation of these systematic effects amounts to 18-34% of the market value of a firm. Furthermore, this systematic heterogeneity is large relative to the mean discount: the standard deviation of systematic heterogeneity attributable to diversification is roughly twice as large as the mean discount. Our estimates imply that, despite an average discount, roughly one-third of firms display a systematic diversification premium. Last, the systematic heterogeneity from diversification is comparable in magnitude to systematic heterogeneity from all other sources combined.

[6] In contrast to these studies, our approach separates between genuine systematic heterogeneity and transitory shocks, while allowing for a flexible pattern of persistence in these shocks.

Our results are robust across the two separate time periods (and datasets) that we study. They are also robust to the inclusion of standard firm-level observables that are typically employed in the prior literature, to outliers, and to controlling for the endogeneity of firm diversification decisions.

In section 2 we illustrate our approach to studying systematic heterogeneity in the context of a simple model. Section 3 describes the data and provides descriptive statistics. Section 4 describes the main empirical specification and certain estimation issues. Section 5 presents the results and describes various robustness checks. Section 6 concludes.

2. The Model

Our approach to studying systematic heterogeneity can be illustrated in the context of a simple model and is motivated by Table 1, which displays the distribution of excess values of diversified and single-segment firms.[7] This table shows that the variation within each of these two groups ("diversified" and "single-segment" firms) swamps the between-group difference in means. For example, the standard deviation of excess values for diversified firms is 4.5-7.9 times as large as the between-group difference in means. Of course, a high cross-sectional variance may simply reflect idiosyncratic shocks that lie largely outside a firm's control. The relevant question is how much of this variance reflects systematic heterogeneity, and what is the type of this systematic heterogeneity (i.e., systematic heterogeneity arising from the diversification choice itself vis-a-vis systematic heterogeneity from other sources).

To illustrate this, we use the following statistical model of excess values for firm i in year t

$$Y_{i,t} = \lambda_t + \alpha_i + (\mu_D + \mu_i)D_{i,t} + u_{i,t} \tag{2.1}$$

where λ_t are the year effects, $\alpha_i \sim N(0, \sigma_\alpha^2)$ is the systematic firm effect, $D_{i,t}$ is a dummy variable equal to 1 if firm i is diversified beyond a single segment in year t, μ_D is the mean

[7] While the term "single-segment" is the common nomenclature in the literature, these firms are occasionally referred to as "specialized", "focused", or "standalone" firms.

diversification discount or premium ($\mu_D < 0$ or $\mu_D > 0$ respectively), $\mu_i \sim N(0, \sigma_\mu^2)$ is the firm-specific systematic diversification effect (additional systematic discount or premium beyond μ_D), and $u_{i,t} \sim N(0, \sigma_u^2)$ represents transitory (possibly persistent, but not permanent) random shocks modeled as an autoregressive process. This model breaks down overall variation in excess values into two parts: transitory shocks ($u_{i,t}$) and systematic heterogeneity (α_i and μ_i). It further breaks down the systematic heterogeneity component into two parts: systematic heterogeneity attributable to the diversification choice (μ_i) versus systematic heterogeneity from all other sources (α_i).

The starting point in modern strategy theory is that imitation gradually erodes most (but not all) sources of superior performance, i.e., they will not be systematic (sustainable). As a result, the primary goal (or raison d'etre even) of strategy is to identify the (relatively rare) sources of superior performance that can be sustained in the long run despite imitation. In terms of our model, such sustainable performance differences based on barriers to imitation will be captured by the systematic heterogeneity terms α_i and μ_i. In particular, μ_i will capture systematic performance differences that are attributable to diversification choices - for example, due to the transfer of unique resources or capabilities across business segments, or the sharing of activities that is hard to imitate (Porter 1996, Collis and Montgomery 2005). α_i will capture systematic performance differences from all other sources except diversification - for example, first mover advantages or scale economies within lines of business, or specific capabilities such as product-specific knowledge in each segment.[8] As is clear, the distinction between μ_i and α_i is important not just for empirical reasons, but theoretical ones too.

In contrast to sustainable (systematic) performance that relies on barriers to imitation, imitable (and therefore unsustainable) sources of superior performance will be captured by the transitory shocks $u_{i,t}$. Notice that these shocks may be somewhat persistent if superior performance from imitable sources erodes gradually, but they are not permanent or systematic. In addition to imitable sources of performance, the transitory shocks $u_{i,t}$ will

[8] As result, and importantly, we allow for a firm to have high α_i but low μ_i (if, for example, these line-of-business specific advantages do not transfer easily across segments), and the reverse.

also capture the effect of external shocks such as fluctuations in demand or input prices. Such external shocks may be somewhat persistent by their nature, also contributing to the persistence in $u_{i,t}$.[9]

The systematic effects of diversification are captured through μ_D and μ_i. The mean diversification discount or premium μ_D captures two things. First, it captures the average causal effect of diversification, i.e., the average value that would be created or destroyed by diversification for an average firm. For example, diversification may systematically destroy value for an average firm by allowing inefficient cross-subsidization of failing business segments (Meyer et al. 1992) or misallocation of capital to less efficient segments due to internal power struggles within the firm (Rajan et al. 2000); or it may systematically create value by creating a more efficient internal capital market (Stulz 1990) or labor market, or allowing it to gain access to unique non-contractible resources from other business segments. Second, empirical findings of a non-zero μ_D may simply reflect econometric issues of endogeneity and self-selection (Campa and Kedia 2002, Villalonga 2004a), as opposed to the average causal effect of diversification. We discuss these econometric issues later on in section 4, and we explicitly control for endogeneity of diversification in robustness checks.

While μ_D captures the value consequences of diversification for an average firm, the systematic heterogeneity term μ_i represents the firm-specific systematic diversification effects beyond the average discount or premium μ_D. In particular, if a firm has a large positive μ_i, it means that this firm has an unusually high ability to systematically create value *through diversification*, and this superior value can be sustained over time. For such a firm, diversification would be a central part of its overall strategy, not just a generic choice of whether to expand or not.[10] At the same time, if a firm has a large negative μ_i,

[9] Even though we interpret $u_{i,t}$ as a persistent but not permanent shock (i.e., a stationary process), our empirical specification of $u_{i,t}$ in estimation does not rule out unit-root processes such as random walk, which correspond to $u_{i,t}$ being a permanent shock. Our empirical estimates indicate that $u_{i,t}$ is in fact stationary, i.e., it only has temporary effect on performance.

[10] For example, diversification is an essential part of Apple's strategy. Thanks to being active in multiple industries (hardware, software, services and retail), with careful integration across products, it is able to offer entire well-differentiated platforms (as opposed to stand-alone products) that are far harder for others to imitate fully.

it means that this firm has an unusually low ability to create value through diversification. This would be the case if, for example, diversification results in negative synergies, creates conflicts of interest, or compromises entrepreneurial autonomy. For such a firm, staying focused (single-segment) would be a necessary part of a successful strategy.

While the prior literature on diversification has focused almost exclusively on the mean discount or premium μ_D (which, as we discuss above, aims to capture the value consequences of diversification for an average firm), we show below that the implications of this mean discount or premium depend crucially on how the overall heterogeneity in excess values decomposes into α_i, μ_i and $u_{i,t}$.

Consider three scenarios. In all of them, we fix the mean discount μ_D at $-\frac{1}{5}$ and the cross-sectional heterogeneity in excess values for diversified firms, $var(\alpha_i + \mu_i + u_{i,t})$, at 1 (these relative magnitudes are consistent with the actual patterns in the data).[11] In other words, in all the scenarios that follow, the cross-sectional heterogeneity in excess values swamps the mean discount. The only difference between these scenarios is in the decomposition of this heterogeneity.

First, suppose that $\sigma_\alpha^2 = \sigma_\mu^2 = 0$, $\sigma_u^2 = 1$, so the heterogeneity in excess values is driven entirely by the transitory shocks. In this case, the mean diversification discount μ_D is the most useful systematic predictor of excess values even though its magnitude is quite small compared to the heterogeneity, and refocusing would be a sensible strategy for all diversified firms. Furthermore, $\sigma_\mu^2 = 0$ means that diversification decisions are "generic" choices that depend little on (fit with) a firm's overall strategy, i.e., the specifics of the diversification decision do not have any systematic value consequences, and therefore do not matter in the long run. Likewise, $\sigma_\alpha^2 = 0$ means that specific attributes of a firm's overall strategy do not have any systematic value consequences and therefore do not matter in the long run (this would arise, for example, if firms only possess generic capabilities or easily imitable resources, or undertake actions that can be quickly mimicked).

Second, consider a scenario where $\sigma_\alpha^2 = 0.5$, $\sigma_\mu^2 = 0$, $\sigma_u^2 = 0.5$, so the heterogeneity is

[11] For the purposes of this illustration, we treat μ_D and μ_i as the *causal* effects of diversification, i.e., we assume that all endogeneity and self-selection issues have been addressed fully in estimation.

driven in part by firm effects, and in part by the transitory shocks. In this hypothetical example, refocusing would still be a sensible strategy for all diversified firms. Like in the first scenario, $\sigma_\mu^2 = 0$ means that diversification decisions are generic, i.e., all firms would systematically destroy (create) value by diversifying as long as $\mu_D < 0$ ($\mu_D > 0$). At the same time, $\sigma_\alpha^2 = 0.5$ means that there are still systematic differences across firms in their ability to create (or destroy) value, but these are not related to the diversification choice itself. *These* systematic drivers, rather than the specifics of the diversification decision, ought to be the focus of managerial consideration in this case.

Third, consider a scenario where $\sigma_\alpha^2 = 0$, $\sigma_\mu^2 = 0.5$, $\sigma_u^2 = 0.5$, so the heterogeneity is driven in part by the firm-specific systematic diversification effects μ_i and in part by the transitory shocks. Even though the mean discount and the cross-sectional heterogeneity in excess values for diversified firms are exactly the same as in the first scenario, the implications now are strikingly different. In this hypothetical example, $\mu_D + \mu_i$ is now positive for 39% of diversified firms.[12] For these firms, diversification would systematically create value *despite* the average discount. Furthermore, $\sigma_\mu^2 = 0.5$ means that while certain firms systematically create value through diversification, others systematically destroy value - hence, understanding the specifics of firms' diversification decisions and their value consequences is crucial. At the same time, $\sigma_\alpha^2 = 0$ means that there are no systematic differences across firms in their ability to create or destroy value other than through diversification. Diversification decisions are, in this case, the decisive drivers of systematic differences in value across firms.

The central insight from these scenarios is that if the systematic heterogeneity in firm-specific diversification effects μ_i is large, then the finding of an average diversification discount or premium would have much less predictive and practical value. At the same time, the μ_i-s can have both predictive and practical value even if the mean discount μ_D is

[12] This percentage is computed as $\Pr(\mu_D + \mu_i > 0)$, where $\mu_D + \mu_i$ is a normally-distributed random variable with mean -0.2 and variance 0.5.

zero, i.e., even if diversification neither creates not destroys value for an *average* firm.[13,14] We therefore turn to the data to decompose overall heterogeneity in values into transitory shocks, systematic firm effects and systematic firm-specific diversification effects.

Last, in addition to the prescriptive implications described above, systematic heterogeneity in the diversification effects μ_i may also have important theoretical implications. To see why, suppose that the estimates of μ_D represents the true average causal effect of diversification (i.e., all endogeneity and self-selection concerns have been addressed properly in estimation). First, consider a world in which there is no systematic heterogeneity in μ_i, i.e., $SD(\mu_i) = 0$. In such a world, findings of a substantial diversification discount or premium μ_D would either be puzzling[15] or reflect irrationality by firms or investors, since they would mean either that many firms persist in systematically value-destroying behavior or investors have systematically wrong expectations.[16] On the other hand, in a world with substantial systematic heterogeneity in the diversification effects μ_i, findings of a substantial average diversification discount or premium μ_D need not reflect either irrationality or inefficiency. Instead, this may simply mean that diversified firms have $\mu_D + \mu_i > 0$, and

[13] In this case, the conclusion based on the mean discount μ_D (the focus of prior literature) would be that diversification does not matter. This, of course, would not be accurate if systematic heterogeneity in μ_i were large, since diversification would matter a lot for a large group of firms – those in both tails of the distribution of μ_i.

[14] Notice that large systematic heterogeneity does not imply a naive prescription that unsuccessful and average firms should imitate successful diversified firms (those with a large positive μ_i). There are at least two reasons why. First, most of the sources of systematic superior μ_i-s are likely inimitable, since this is the primary mechanism that makes μ_i-s systematic in the first place. Second, even if perfect imitation of a successful diversified firm is possible, it is not necessarily desirable if it results in ruinous head-on competition. While systematic heterogeneity does not give managers exact recipes on how to achieve sustainable superior performance, it provides useful guidance on where they should focus most of their attention in an effort to achieve sustainable superior performance.

[15] Indeed, several papers have tried to offer rational interpretations of the discount "puzzle", for example based on real options (Bernardo and Chowdhry 2002), agency (Morck, Shleifer, and Vishny 1990), lower expected returns (Lamont and Polk 2001), or search for new profit opportunities after the firm has become relatively unproductive in its current activities (Gomes and Livdan 2004), among other reasons. Also, Levinthal and Wu (2010) show that optimal diversification that increases firm profit may at the same time reduce the profit margin and market-to-book ratio (common metrics for the consequences of diversification) due to the spread of non-scale free capabilities across additional segments. Adner and Zemsky (2006) point out that rational diversification decisions may affect the equilibrium market structure (and therefore profitability), which also may generate spurious findings of a discount.

[16] In such a world, an efficient outcome would involve all firms being single-segment (if $\mu_D < 0$), or all firms being diversified (if $\mu_D > 0$). In either of these efficient outcomes, μ_D would be invisible to the researcher since there would not be any cross-sectional or time-series variation in the diversification status in the data.

single-segment firms have $\mu_D + \mu_i < 0$, with all firms behaving optimally.[17]

3. Data and Descriptive Statistics

Our estimation samples are based on Compustat firm-level and segment-level data for two subperiods: 1978-1996 and 1998-2008. Due to a change in segment reporting requirements in 1997 (adoption of SFAS 131—see Berger and Hann 2003 for details), the pre-1997 and post-1997 segment-level data are not directly comparable with each other. Therefore, we discard the transition year 1997, and we analyze each subsample separately. The 1978-1996 sample corresponds to the longest time period for which segment-level data are available and reported in a uniform format (SFAS 14). It is similar to the data used in Campa and Kedia (2002) and Villalonga (2004a), which facilitates the comparison of our findings to the prior literature on diversification. The 1998-2008 sample extends the analysis to a more recent period, which is of particular interest because companies' diversification behavior may have changed since the 1990s as the notion of the diversification discount gained influence.

Our sample selection criteria follow Berger and Ofek (1995) and Villalonga (2004a). We start from the full sample of firms that appear in both the firm-level and the segment-level Compustat files between 1978-1996 or 1998-2008. We discard firm-years if firm-level sales are below \$20 million or missing, firm-level assets or market value are missing or negative, the sum of segment sales differs from firm-level sales by more than 1%, or the sum of segment assets differs from firm-level assets by more than 25%.[18] We also discard firm-years if segment sales or assets are missing or negative for any of the segments, if SIC code is missing for any of the segments, if any of the segments is in the financial sector, agriculture, government and other noneconomic activities (one-digit SIC code 0, 6 or 9),

[17] For example, diversification may still destroy value on average by leading to inefficient cross-subsidization for all firms. If so, the diversified firms would be the ones for which the value loss from cross-subsidization is less than the unique strategic benefits they can generate through diversification.

[18] Unlike sales, many assets cannot be meaningfully traced to a specific segment (e.g., corporate headquarters assets), therefore Berger and Ofek (1995) use the less stringent 25% threshold for assets. Following Berger and Ofek, if the sum of segment assets differs from firm-level assets by less than 25%, we rescale the segment assets so that they will add up to the firm-level total assets.

or if any of the segments has less than 5 single-segment firms in its 2-digit SIC.[19] Since our estimation model relies on the time-series dimension of the data, we also discard firms that have only one valid firm-year in the data.[20] These criteria yield the final samples of 52,803 firm-years for 7,052 firms between 1978-1996, and 29,730 firm-years for 5,412 firms between 1998-2008.

We use three alternative measures of excess values. Following Berger and Ofek (1995), we compute two measures of excess values based on asset and sales multipliers. Both measures are computed as the natural logarithm of the ratio of a firm's actual market value to its imputed value. The market value of a firm is computed as the market value of common equity plus the book value of debt and preferred equity. The imputed value of a firm is the sum of the imputed values of its segments, where a segment's imputed value is equal to the segment assets (sales) multiplied by the median ratio of market value to assets (sales) for the single-segment firms in the corresponding industry-year. The third measure of excess values is based on industry-adjusted q following Lang and Stulz (1994). We compute it as the percentage difference between a firm's actual q and its imputed q.[21] The imputed q for a firm is computed as the asset-weighted average of the imputed q's of its segments, where a segment's imputed q is computed as the average q for the single-segment firms in the corresponding industry-year, and q is measured as the ratio of market to book value. For all three measures of excess values, the industry-year medians and averages are computed at the most precise SIC level for which we observe at least five single-segment firms in the industry-year: 53.3% (58.0%) at the 4-digit SIC level, 26.1% (20.4%) at the 3-digit level and 20.6% (21.6%) at the 2-digit level for the 1978-1996 (1998-2008) sample.[22]

[19] In the post-1997 data, many firms report a reconciliation segment that bridges between segment-level and firm-level disclosures (for example, it may account for unallocated corporate assets and intra-firm sales from one segment to another). Since the reconciliation segment does not have a meaningful industry definition, we allocate its sales and assets among the rest of the segments, and we do not require it to have a valid SIC code.

[20] This screening criterion follows Campa and Kedia (2002). In robustness checks, the estimation results are similar when we keep such firms in the sample, and when we require firms to have at least 5 valid observations, or to have no missing observations throughout the sample period.

[21] While Lang and Stulz (1994) compute their excess value measure as $actual\ q - imputed\ q$, we rescale it to $\frac{actual\ q - imputed\ q}{imputed\ q}$. Since our focus is on variances rather than means, this rescaling is important since it makes the variances more comparable across low-q and high-q industries.

[22] The proportion of matches at different SIC levels for the 1978-1996 sample is similar to Campa and

The descriptive statistics are presented in Table 1. Following Lang and Stulz (1994) and Berger and Ofek (1995), we distinguish between single-segment and diversified (multi-segment) firms. The mean (median) discount in Table 1 is computed as the mean (median) difference in excess values between single-segment and diversified firms. The mean (median) discount in the 1978-1996 sample is 11.1% (10.5%) using asset multipliers, 12.4% (13.0%) using sales multipliers, and 13.7% (10.2%) using industry-adjusted q's, similar to the estimates in the prior literature. The mean (median) discount in the 1998-2008 sample is very similar: 10.1% (9.6%) using asset multipliers, 11.5% (13.5%) using sales multipliers, and 11.9% (8.2%) using industry-adjusted q's. There is also substantial cross-sectional heterogeneity in excess values. For example, for all three measures of excess values in both periods, the standard deviation of excess values for diversified firms is 4.5-7.8 times as large as the median diversification discount, and between 37%-43% of diversified firm-years have higher excess values than the median single-segment firm. However, as we discuss in section 2, the implications of this cross-sectional heterogeneity depend crucially on how it decomposes into transitory (but possibly persistent) shocks $u_{i,t}$, systematic heterogeneity related to diversification μ_i, and systematic heterogeneity from other sources α_i.

Next, as a simple descriptive exercise, we decompose the overall variation in excess values using a fixed-effects approach since most of the prior variance decomposition literature discussed in the introduction used analogous fixed-effects specifications. In particular, we use OLS to decompose the overall variation in excess values into firm fixed effects, mean diversification discount or premium, firm-specific diversification effects (discounts or premiums), and transitory shocks. The estimation equation is the fixed-effects analogue of our main model (2.1)

$$Y_{i,t} = g_t + a_i + (m_D + m_i)D_{i,t} + v_{i,t}$$

where $Y_{i,t}$ is excess value for firm i in year t, g_t is the year effect, a_i is the firm fixed effect (analogous to α_i in our main model), $D_{i,t}$ is the diversification dummy, m_D is the mean diversification discount or premium (analogous to μ_D in our main model), m_i is the

Kedia (2002), who report 50.0% matches at the 4-digit level, 26.5% at the 3-digit level, and 23.5% at the 2-digit level.

firm-specific diversification effect (i.e., an additional fixed effect for the diversified firm-years of firm i, analogous to μ_i in our main model), and $v_{i,t}$ is an error term.[23,24] The cross-sectional variation in m_i and a_i captures heterogeneity in excess values across firms.

Table 2 presents the standard deviations of m_i and a_i across firms, based on the firm-specific estimates of a_i and m_i.[25] These standard deviations provide simple descriptive measures of the magnitude and sources of systematic heterogeneity in excess values. For all three measures of excess values in both samples, the standard deviation of the firm-specific diversification effects m_i is at least 3.4 times as large as the median and mean discounts documented in Table 1, suggesting large systematic variation in the diversification discounts or premiums across firms. Also, the standard deviation of the firm fixed effects a_i is the same order of magnitude as the standard deviation of m_i, suggesting that (1) there is also substantial systematic heterogeneity from other sources besides diversification, but (2) diversification is a very important source of systematic heterogeneity, accounting for a sizable fraction of the total systematic heterogeneity from all sources.

However, the standard deviations of a_i and m_i in this simple fixed-effects approach are likely to significantly overestimate the magnitude of the *systematic* part of overall heterogeneity in excess values, since the fixed effects a_i and m_i pick up not only the actual systematic components of excess values (α_i and μ_i in terms of our main model), but also the firm-specific averages of the transitory shocks $v_{i,t}$.[26] In fact, when we compare these fixed-effects estimates of systematic heterogeneity to our main estimates described in section 5, we find that the fixed-effects approach suffers from a substantial upward bias, overstating the systematic part of overall heterogeneity by over 35% for the diversification effects μ_i,

[23] The notation here is different from our main model (a_i instead of α_i, and m_i instead of μ_i), to emphasize that these are fixed effects as opposed to random effects in our main estimation model.

[24] For consistency with our main model, we normalize the mean of the fixed diversification effects m_i to zero, and estimate their mean using a separate parameter m_D (analogous to μ_D) in our main model.

[25] For firms that remained diversified throughout the sample period, fixed-effects estimation yields estimates of the combined $a_i + m_i$ but not a_i and m_i separately. Therefore we do not include them in our computation of the standard deviations of a_i and m_i. Also, m_i cannot be identified for firms that remained single-segment throughout the sample period, therefore we do not include them in our computation of the standard deviation of m_i.

[26] In other words, the firm fixed effects are unbiased but noisy estimates of the true systematic component of excess values. Consequently, they are appropriate for inferences about means but not about variances.

and by over 40% for the firm effects α_i. This indicates that much of the variation in the fixed effects is actually driven by the transitory shocks, which are possibly persistent but not permanent (systematic). Therefore, it is essential to carefully separate between systematic heterogeneity and transitory shocks in estimation. We do this in the next section using a random-effects approach, which relies on explicitly modeling the variance-covariance structure of the data (including the contribution of persistent but not permanent shocks to this variance-covariance structure) as opposed to inferences based on firm-specific parameter estimates as in the fixed-effects approach.

4. Empirical Model Specification and Estimation Issues

Our main empirical specification follows model (2.1) from section 2. In this section, we add details relevant for estimation, and discuss how we handle potential endogeneity issues. The excess value of firm i in year t follows

$$Y_{i,t} = \lambda_t + \alpha_i + (\mu_D + \mu_i)D_{i,t} + u_{i,t}$$

where λ_t represents the year effects, $D_{i,t}$ is a dummy variable equal to 1 if firm i is diversified (multi-segment) in year t, α_i is the firm effect, μ_D is the mean diversification discount or premium, μ_i is the firm-specific systematic diversification effect (additional discount or premium beyond μ_D), and $u_{i,t}$ represents transitory shocks which may be persistent (but not permanent). In robustness checks, we also estimate various extensions of this main specification.

We specify the systematic heterogeneity terms α_i, μ_i as normally-distributed random effects with mean zero, variances σ_α^2 and σ_μ^2 respectively, and covariance $\sigma_{\alpha\mu}$. As we discuss in section 2, they capture *systematic* performance differences, i.e., performance differences that can be sustained in the long run despite imitation. We specify the transitory shocks $u_{i,t}$ as a flexible autoregressive process with K lags, $u_{i,t} = \sum_{\tau=1}^{K} \rho_\tau u_{i,t-\tau} + \varepsilon_{i,t}$, where $\varepsilon_{i,t} \sim N(0, \sigma_{\varepsilon D}^2)$ for diversified firms, and $\varepsilon_{i,t} \sim N(0, \sigma_{\varepsilon S}^2)$ for single-segment firms. This flexible specification allows us to carefully capture the persistence patterns in $u_{i,t}$. As we discuss in section 2, $u_{i,t}$ captures imitable (and hence unsustainable) sources of superior performance, as well as external shocks.

The empirical specification above assumes that the unobservables α_i and μ_i are independent of the diversification dummy $D_{i,t}$. In other words, the difference in means between diversified and single-segment firms is attributed solely to the mean diversification discount μ_D. However, Campa and Kedia (2002) and Villalonga (2004a) provide strong evidence for endogeneity of the diversification choices. Framed in terms of our empirical model, their findings suggest that $E(\alpha_i + \mu_i | D_{i,t} = 1) < E(\alpha_i + \mu_i | D_{i,t} = 0)$, i.e., diversified firms are a self-selected subsample of firms with inferior unobservables.[27] To the extent that firms' diversification choices are correlated with their α_i and μ_i, the estimate of μ_D in our baseline specification will not have a valid interpretation as the average causal effect of diversification.

We address the endogeneity concerns in several ways. First, we show that, without any control for endogeneity in estimation, our baseline estimates of systematic heterogeneity (the standard deviations of μ_i and α_i) are likely to provide reliable lower bounds for the true magnitude of systematic heterogeneity. At the same time, based on prior literature, our baseline estimates of μ_D are likely to overstate the true mean diversification discount. This, combined with the lower bound for systematic heterogeneity, will also give us a reliable lower bound on how important the systematic heterogeneity is relative to the mean discount. Second, in robustness checks we go beyond this by explicitly controlling for endogeneity and selection using several alternative approaches that mostly follow Campa and Kedia (2002).

Our lower-bound argument for the estimates of systematic heterogeneity is as follows. Suppose that the propensity to diversify is increasing in μ_i. In other words, firms that expect to gain more from diversification are more likely to diversify (or to stay diversified), and firms that expect to lose more from diversification are more likely to refocus (or to stay focused).[28] If so, we will have four groups of firms in the data that self-select

[27] The empirical analysis in Campa and Kedia (2002) and Villalonga (2004a) does not separate between α_i and μ_i, and focuses on the combined unobservable ($\alpha_i + \mu_i$ in terms of our model).

[28] This assumption does not contradict the findings of Campa and Kedia (2002) and Villalonga (2004a).

based on their μ_i: (1) firms that remained diversified throughout the sample period, (2) firms that diversified during the sample period, (3) firms that refocused during the sample period, and (4) firms that remained single-segment throughout the sample period. Since the propensity to diversify is increasing in μ_i, the first group (diversified firms) is likely to have disproportionately high μ_i-s, the two groups in the middle (diversifying and refocusing firms) are likely to have moderate μ_i-s, and the fourth group (single-segment firms) is likely to have disproportionately low μ_i-s. The standard deviation of μ_i is identified by the first three groups (diversified firms, diversifying firms, and refocusing firms),[29] which represent the upper tail and the moderate part of the distribution of μ_i-s. At the same time, the fourth group (single-segment firms), which disproportionately represents the lower tail in the distribution of μ_i, plays no role in identifying the standard deviation of μ_i because their excess values do not contain μ_i in any way. Thus, the estimates of systematic heterogeneity in μ_i will be based on three self-selected groups that have a more homogeneous distribution of μ_i-s than the entire sample. Consequently, the estimates of $SD(\mu_i)$ based on these three self-selected groups will provide a lower bound for the true systematic heterogeneity in μ_i for the entire sample.

In contrast to μ_i, the standard deviation of α_i is identified based on all firms in the full sample. It is easy to show that, for any self-selection rule with respect to α_i, the estimates of $SD(\alpha_i)$ based on self-selected groups of firms will be lower than the true standard deviation of α_i for the full sample. Specifically, self-selection into diversification based on α_i implies that the mean α_i for the diversified firm-years[30] will be different from the mean α_i for the single-segment firm-years. In estimation, the mean discount or premium μ_D

Their findings suggest that diversified firms have lower combined unobservable $\alpha_i + \mu_i$, but they do not provide any evidence on μ_i alone. By combining their findings with the simple argument that firms that gain from diversification are more likely to diversify, we expect the diversified firms to be predominantly the low-α_i high-μ_i firms.

[29] The two groups in the middle (diversifying and refocusing firms) provide a direct source of identification for $SD(\mu_i)$, since the systematic part of their excess values is based on α_i in their single-segment years, and $\alpha_i + \mu_i$ in their diversified years. The diversified group provides an additional source of identification for $SD(\mu_i)$, however it is less direct because the systematic part of their excess values is based on the combined $\mu_i + \alpha_i$ in all years.

[30] By "diversified firm-years", we mean all years for diversified firms and diversified firm-years for firms that changed their diversification status.

will capture the difference in mean α_i between diversified and single-segment firm-years (the between-group variation), and the estimate of $SD(\alpha_i)$ will only capture the residual (within-group) variation in α_i around its means within the respective self-selected groups. Since the residual variation in α_i around its group-specific means is less than its total variation, the estimates of $SD(\alpha_i)$ will provide a lower bound for the true variation in α_i for the entire sample.[31]

Thus, our baseline estimates, which do not control for endogeneity, will yield a lower bound for the true magnitude of systematic heterogeneity in α_i and μ_i. At the same time, the findings of Campa and Kedia (2002) and Villalonga (2004a) indicate that our estimate of μ_D will likely overstate the true mean discount for our 1978-1996 subsample (which is similar to their data), and, to the extent that the same patterns persist in more recent years, we will have a similar bias in μ_D in our 1998-2008 subsample.[32] Thus, by combining an overstated mean discount with the lower bound for systematic heterogeneity, we will also obtain a reliable lower bound for how important the systematic heterogeneity is relative to the mean diversification discount.

In robustness checks, we go beyond this argument and explicitly control for endogeneity of diversification choices. However, one important advantage of our lower-bounds approach is that it allows us to get reliable lower bounds on the magnitude of systematic heterogeneity without imposing strong additional assumptions.

5. Estimation Results

We estimate the model using maximum likelihood. The optimal number of lags for the autoregressive shock $u_{i,t}$ is chosen based on the Bayes Information Criterion (BIC). We estimate the model for each of the three excess values measures (based on asset multipliers and sales multipliers following Berger and Ofek 1995, and based on industry-adjusted q-s

[31] Since total variation in α_i for the entire sample is the sum of between-group variation and within-group variation, the within-group variation is always (weakly) less than the total.

[32] When we explicitly control for the endogeneity of diversification in robustness checks, the direction of change in μ_D in both samples is in fact consistent with our prior expectations based on Campa and Kedia (2002) and Villalonga (2004a).

following Lang and Stulz 1994), for the 1978-1996 and 1998-2008 samples.

Our baseline estimates are presented in Table 3.[33] The main parameter of interest is $SD(\mu_i)$, which measures the systematic heterogeneity in the firm-specific diversification effects (discounts or premiums). The estimates of $SD(\mu_i)$ range from 0.18 to 0.34 for both time periods and for all three excess value measures, indicating very substantial heterogeneity across firms in the systematic effects of diversification.[34] For example, if we compare a firm whose μ_i is one standard deviation above the mean against a firm with mean μ_i, the difference between them in the value from diversification amounts to 18-34% of the total market value of the firm. Furthermore, this gap is systematic, i.e., it will persist year after year. This indicates that diversification is a very important strategic variable in the sense that it represents a major source of sustainable superior performance for many firms.

The systematic heterogeneity in the diversification discount is also large relative to the mean discount μ_D. For all three measures of excess values in both time periods, the standard deviation of μ_i is 1.7-2.7 times as large as the mean discount. Framed differently, our estimates imply that the mean diversification discount (μ_D) accounts for only 9-19% of total systematic variation in excess values due to diversification, while other firm-specific systematic factors related to diversification (μ_i) account for the remaining 81-91%.[35] The estimates indicate that, despite the average discount, 28-35% of firms actually display a systematic diversification premium ($\mu_D + \mu_i > 0$). If we treat μ_D as a

[33] For brevity, we do not report t-values or significance levels in Table 3. All parameters of interest in Table 3 are significant at the 0.1% level or better.

[34] Since our excess values measures are computed as $ln\left(\frac{\text{market value}}{\text{imputed value}}\right)$, the units of $SD(\mu_i)$ are directly interpretable as percentages of market value.

[35] These percentages are based on the following calculation. The systematic effect of diversification on excess value of firm i is $D_{i,t}(\mu_D + \mu_i)$, where $D_{i,t}$ is the diversification dummy. Variation in $D_{i,t}$ and μ_i leads to systematic variation in excess values across firms (both between diversified and single-segment firms, and within the group of diversified firms) that is due to these systematic effects of diversification. The total contribution of diversification to systematic variation in excess values is $var(D_{i,t}(\mu_D+\mu_i))$, which can be decomposed as

$$var(D_{i,t}(\mu_D+\mu_i)) = cov(D_{i,t}(\mu_D+\mu_i), D_{i,t}(\mu_D+\mu_i)) = cov(D_{i,t}\mu_D, D_{i,t}(\mu_D+\mu_i))+cov(\mu_i, D_{i,t}(\mu_D+\mu_i))$$

The first term in this decomposition captures the relative contribution of μ_D, and the second term the relative contribution of μ_i.

valid estimate of the average causal effect of diversification, it indicates that, even though diversification would systematically destroy value for an average firm, 28-35% of firms can actually systematically create value by diversifying. As we discuss earlier, our baseline estimates likely overstate the mean discount and understate the true systematic heterogeneity. Consequently, the actual proportion of firms that would systematically create value by diversifying is likely to be even higher.[36]

In addition to large systematic heterogeneity related to diversification (μ_i), we also find large systematic heterogeneity from all other sources (α_i). However, the magnitude of $SD(\mu_i)$ is comparable to $SD(\alpha_i)$, i.e., systematic heterogeneity related to diversification accounts for a large fraction of the overall systematic heterogeneity from all sources. When we look at the variance of total systematic heterogeneity for diversified firms, $\mu_i + \alpha_i$, we find that μ_i accounts for 21% to 63% of this total systematic heterogeneity.[37] This further reaffirms that diversification is a very important strategic variable. However, it is important primarily due to systematic heterogeneity—i.e., as a source of systematic performance differences related to the firm-specific diversification effects—rather than because of the mean effect μ_D that has been the focus of the prior literature.

Our baseline estimates in Table 3 carefully separate between systematic heterogeneity (α_i and μ_i) and transitory (possibly persistent but not permanent) shocks $u_{i,t}$. One instructive comparison is between our estimates of systematic heterogeneity and the simple OLS fixed-effects estimates reported in Table 2, which follow the approach used in most of the prior literature on variance decomposition of performance. The comparison is striking: the fixed-effects estimates in Table 2 overstate the systematic heterogeneity related to diversification by over 35%, and they overstate the systematic heterogeneity from other sources by over 40%.

[36] It is indeed higher in the robustness checks where we explicitly control for endogeneity.

[37] This decomposition is as follows. We can rewrite $var(\alpha_i + \mu_i) = cov(\alpha_i, \alpha_i + \mu_i) + cov(\mu_i, \alpha_i + \mu_i)$. Therefore, the ratios $\frac{cov(\alpha_i,\alpha_i+\mu_i)}{var(\alpha_i+\mu_i)}$ and $\frac{cov(\mu_i,\alpha_i+\mu_i)}{var(\alpha_i+\mu_i)}$ add up to 100%, and they can be interpreted as the relative contribution of α_i and μ_i respectively to the total variance of $\alpha_i + \mu_i$.

5.1. Robustness Checks

5.1.1. Observables

As a first robustness check, we re-estimate the model after adding control variables from Berger and Ofek (1995): log total assets (a proxy for size, which is positively correlated with diversification), EBIT/sales (profitability), CAPEX/sales (a proxy for growth opportunities), and related diversification.[38] This robustness check is essential for two reasons. First, the systematic heterogeneity in our baseline estimates in Table 3 could simply reflect variation explained by standard control variables identified in the prior literature. Second, if these simple observable variables explained a large fraction of systematic heterogeneity, one may even question the interpretation that such heterogeneity represents performance differences that are hard to imitate.[39] The estimation results after adding Berger-Ofek controls are presented in Table 4. While these control variables are highly significant, the estimates of systematic heterogeneity in α_i and μ_i are almost identical to our baseline estimates in Table 3. In another robustness check, we use the control variables from Lang and Stulz (1994): log total assets (a proxy for size), R&D intensity, and a dummy for payment of dividends.[40] The results are very similar (untabled). And, in another robustness check, we control for the number of segments for diversified firms in order to capture the degree of diversification (which may be one of the main factors behind the systematic heterogeneity in μ_i). As expected, the mean diversification discount is larger (in absolute value) for diversified firms with more segments. Again, however, the estimates of systematic heterogeneity in α_i and μ_i are virtually identical to our baseline estimates. These results suggest

[38] We compute the related diversification variable as follows. First, following Berger and Ofek (1995), for each firm-year we aggregate segments that have the same 2-digit SIC. After that, for each firm-year, we compute the Herfindahl index of its sales across the 2-digit SICs to characterize how dispersed its segment sales are across different 2-digit industries.

[39] Strategy theory suggests that the root drivers of sustainable performance differences are unlikely to be summarized by a few generic variables. Therefore, while we expect the Berger-Ofek controls to have some explanatory power, we do not expect them to account for a large fraction of systematic heterogeneity.

[40] Lang and Stulz (1994) control for R&D intensity because, unlike capital investments, investments in R&D are not capitalized (recognized as an asset on the balance sheet), which mechanically increases the market-to-book ratio. They control for payment of dividends because cash-constrained firms that do not pay dividends may not be able to raise enough capital because of capital market imperfections, in which case they will not be able to exhaust all their positive-NPV projects, and consequently their marginal q will be above one.

the systematic heterogeneity estimates captured here are unlikely to be traced to simple observables.

In another robustness check, we re-estimate the model after adding commonly-used proxies for firm strategy (e.g., Kotha and Nair 1995, Spanos et al. 2004): industry-adjusted R&D intensity and advertising intensity as proxies for differentiation, market share and log total assets as proxies for economies of scale or scope (one dimension of cost efficiency), assets/sales and CAPEX/sales as measures of "asset parsimony" (another dimension of cost efficiency following Kotha and Nair 1995), and the same variables squared to capture potential non-linear effects of "stuck in the middle."[41] The strategic proxies are highly significant, however the magnitude of systematic heterogeneity (Table 5) is again very similar to our baseline estimates. Again, these results reinforce the interpretations from aforementioned theories of strategic advantage— that systematic heterogeneity captures sustainable performance differences which are unlikely to be meaningfully reduced to a small number of generic strategic proxies.

5.1.2. Outliers

To ensure that the estimates are not driven by a small number of extreme observations, we also re-estimate all models after discarding 1% outliers on each tail for excess values. The estimates are similar (untabled).[42] We also tried varying the threshold for the minimum number of valid firm-years a firm should have in order to be included in our estimation sample (our main sample includes all firms that have at least 2 valid firm-years in order to maximize the usable time-series information in the data). The results are robust to adding firms that had only a single valid firm-year, restricting the sample to firms that

[41] We compute these "industry-adjusted" variables as deviations of the original variables from the imputed industry medians for each year. For single-segment firms, the imputed industry median is simply the median of the respective strategic variable for all single-segment firms within its industry-year. For diversified firm, the imputed industry median is computed as the sales-weighted average of industry medians for its segments. This procedure is similar to that used in computing imputed firm values based on sales multipliers. Compared to using the original variables, the use of industry-adjusted variables improves the likelihood substantially, while the estimates of systematic heterogeneity are very similar in both specifications.

[42] In general, discarding outliers should reduce the estimates of variances to below their true value. As expected, the estimates of variances are slightly (but not dramatically) lower.

had at least 5 valid firm-years, and restricting the sample to firms that had no missing observations throughout the sample period.

5.1.3. Endogeneity of Diversification Choice

As we discuss in section 4, our baseline estimates of the mean diversification discount μ_D are likely to overestimate its true magnitude due to endogeneity of the diversification choice. Therefore, in the following robustness checks we explicitly control for endogeneity using several alternative methods.

In one robustness check, we use the random-effects analogue of the fixed-effects approach used in Campa and Kedia (2002). In this approach, the firm fixed effects pick up the unobservables that remain constant over time (the equivalent of α_i in our model), and the mean diversification discount is identified longitudinally based on firms that changed their diversification status during the sample period. As we show earlier in this section, fixed-effects estimation is not appropriate for estimating the variances of systematic heterogeneity (however, it is fully appropriate for estimating the mean discount which is the focus of Campa and Kedia 2002). Therefore, rather than estimate a fixed-effects model, we use an analogous random-effects approach: we estimate our baseline model only for the subsample of firms that diversified or refocused during the sample period (the "switchers"). The estimation results for the "switchers" are presented in Table 6. As expected, for all three measures of excess values in both time periods, the estimates of the mean discount μ_D are substantially closer to zero than our baseline estimates in Table 3. This is consistent with the fixed-effects estimates in Campa and Kedia (2002) and the longitudinal estimates in Villalonga (2004a). The direction of change in μ_D relative to our baseline estimates for the full sample suggests that diversification is negatively correlated with the firm effect α_i, i.e., firms that have systematically poorer performance on dimensions unrelated to diversification are more likely to diversify. At the same time, the estimates of systematic heterogeneity (the standard deviations of α_i and μ_i) are similar to the baseline estimates. Since the estimates of the mean discount μ_D are now closer to zero, the proportion of firms enjoying a systematic diversification premium ($\mu_D + \mu_i > 0$) is now estimated to be between 34-41%, higher than in the baseline estimates.

While this longitudinal approach controls for endogeneity that arises due to correlation between the firm effects α_i and the diversification choices, it may be vulnerable to correlation between the transitory shocks $u_{i,t}$ and diversification (for example, a firm may seek to diversify after its performance dropped due to adverse $u_{i,t}$-s). As an additional robustness check to address this concern, we modify the model for "switchers" by adding dummies for each year from two years before to two years after the diversification or refocusing event (we estimate two separate sets of parameters on these dummies for diversifying and refocusing firms).[43] We also add parameters to allow for a different variance-covariance structure of the transitory shocks $u_{i,t}$ around the diversification or refocusing event. The estimates in this modified model are very similar (untabled).

In another robustness check, we control for self-selection using the Heckman selection model following Campa and Kedia (2002). We employ their instruments for the diversification dummy $D_{i,t}$: the proportion of other diversified firms in the 2-digit industry of firm i in year t, the proportion of other firms' sales in the 2-digit industry of firm i in year t that is due to other diversified firms, S&P index dummy, a major stock exchange dummy and a foreign incorporation dummy.[44] In the first stage, we run a probit regression of the diversification dummy on the instruments and control variables: log total assets for years t, $t-1$, $t-2$, EBIT/sales for years t, $t-1$, $t-2$, CAPEX/sales for years t, $t-1$, $t-2$, leverage, and log total assets squared (the control variables follow Campa and Kedia 2002). In the second stage, we re-estimate our baseline model after adding the inverse Mill's ratio and the control variables from the first stage. The estimation results are presented in Table 7. As expected, the correction for self-selection changes the estimates of the mean discount μ_D dramatically. For all three excess values measures in both time periods, the estimates of μ_D change from a substantial discount in the baseline estimates to either a negligible discount or a diversification premium. In all cases, the coefficient on the inverse

[43] The Heckman selection correction in the next paragraph also addresses this concern.

[44] Campa and Kedia use several additional instruments that vary over time but not across firms (number of M&A announcements, M&A volume in dollars, and macroeconomic variables). Since our main model includes year effects, we also include year effects in the first-stage probit model for consistency. Due to the inclusion of year effects, these instruments drop out.

Mill's ratio is negative and significant at the 5% level, indicating negative correlation between self-selection into diversification and the combined unobservable $\alpha_i + \mu_i + u_{i,t}$.[45] At the same time, the estimates of systematic heterogeneity are very close to our baseline estimates, indicating that our main results are robust to controlling for self-selection.

5.2. Generalized versus Diversification-related Systematic Heterogeneity: Consequences of Misspecification

Central to our approach is distinguishing between "generic" systematic heterogeneity—for example, capturing differences in firm-level resources such as knowledge, R&D skills, patents, brand strength – and systematic heterogeneity related to a firm's diversification choice (i.e., α_i versus μ_i). In contrast, studies that aim to capture performance differences across firms typically do not distinguish between the two. Most common are models that estimate mean discount or premium and employ "firm fixed effects" to capture everything that is systematically different across firms (i.e., α_i and μ_i are lumped together for diversified firms).

In this subsection, we illustrate the biases in firm-level inferences that arise from lumping together α_i and μ_i into a single "firm effect" $\widetilde{\alpha}_i$. To do that, we use the full sample of firms that changed their diversification status between 1978-1996, and estimate two competing models: (1) the "restricted" version of our main model, in which we model overall systematic heterogeneity using only firm effects $\widetilde{\alpha}_i$, i.e., we restrict all μ_i-s to be zero following the standard approach, and (2) the "full" model, in which we separately model "generic" systematic heterogeneity α_i and systematic heterogeneity related to diversification μ_i. Notice that even though $\widetilde{\alpha}_i$ and α_i enter the respective model similarly, they have different interpretations: $\widetilde{\alpha}_i$ in the restricted model captures overall systematic heterogeneity from all sources, while α_i in the full model only captures "generic" systematic heterogeneity that is unrelated to diversification. After estimating these models, we compute the posterior mean of $\widetilde{\alpha}_i$ in the restricted model, and the posterior means of α_i

[45] Since the systematic diversification effect μ_i enters excess values only for diversified firms, we also re-estimate the model allowing for different coefficients on the inverse Mill's ratio for diversified and single-segment firms. The results are very similar.

and μ_i in the full model, for each firm i in the sample. These posterior means represent the best firm-level estimates of the systematic components of excess values for firm i, based on the time-series excess values data for firm i and based on the structure of the model.[46] We use the posterior means from the two models to examine to what extent the firm effects in the standard approach (the restricted model) are able to capture "generic" systematic heterogeneity versus systematic heterogeneity related to diversification.

Figure 1 plots the posterior means of the firm effect $\widetilde{\alpha}_i$ in the restricted model against the posterior means of α_i in the full model (to keep the graph manageable and to be able to highlight specific examples of firms, we only plot the 100 largest firms by assets[47]). Overall, the firm effects in the restricted model are fairly close to the estimates of "generic" systematic heterogeneity in the full model (the correlation between them is 0.95). However, there are large differences for some firms. For example, the restricted model underestimates the "generic" part of systematic heterogeneity in excess values by 0.22 (equivalent to 22% of market value) for 7-Eleven, by 0.16 for Coca-Cola, and by 0.14 for Walmart, and overestimates it by 0.06 for Komatsu, by 0.09 for Macy's and by 0.22 for Ceridian Corp. In other words, the firm effects in the standard approach underestimate how successful 7-Eleven, Walmart, or Coca Cola is on dimensions unrelated to diversification, and they overestimate it for Komatsu, Macy's or Ceridian.

Figure 2 plots the posterior means of the firm effects $\widetilde{\alpha}_i$ in the restricted model against the posterior means of the firm-specific diversification effect μ_i. The correlation between them is -0.11, i.e., even though the firm effects in the standard approach successfully identify firms that are systematically successful on dimensions unrelated to diversification (α_i), they fail to identify firms that are systematically successful or systematically unsuccessful in their diversification choices. For example, while Walmart and Coca-Cola have the

[46] Technically, both OLS fixed effects and our posterior means represent weighted averages of the time-series data for firm i. However, in contrast to the fixed effects, the posterior means in our models recognize that a large part of the variation in excess values is due to (somewhat persistent) transitory shocks, and adjust the weights on the data accordingly. As expected, when we estimate the fixed-effects analogues of our resticted and full models, the standard deviation of the fixed effects turns out to be more than twice as large as the standard deviations of the posterior means. This suggests that the fixed-effects estimates in fact misinterpret transitory shocks, and wrongly attribute them to systematic differences across firms.

[47] The patterns for the full sample used in estimation are similar.

highest firm effects among the 100 large firms plotted in Figures 1 and 2, they have disproportionately low estimates of the diversification effects μ_i. In other words, while these are very successful firms overall (based on the firm effects $\widetilde{\alpha}_i$, or based on the total $\alpha_i + \mu_i$ in the full model), they are not estimated to be successful diversifiers.[48] At the same time, while Komatsu and Macy's are approximately average overall (based on their firm effects), they are among the more successful diversifiers (based on μ_i). Similarly, while Ceridian is below-average overall (based on its firm effect), it is a highly-successful diversifier (i.e., it has a high μ_i combined with a below-average α_i). These estimates are consistent with casual observations regarding these firms. For example, Walmart and Coca-Cola are traditional textbook examples of very successful firms (consistent with their high α_i). At the same time, they are not typically referred to as noteworthy diversifiers (consistent with their low μ_i in our estimates).

These sharp differences between "generic" systematic heterogeneity α_i and systematic heterogeneity related to diversification μ_i reinforce the importance of carefully separating the different types of heterogeneity when investigating performance differences across firms.

6. Conclusion

Understanding value creation from diversification is one of the central questions in the study of multibusiness strategy and corporate finance. In contrast to prior empirical work on this topic, whose dominant focus is the average discount or premium for multibusiness firms, this paper sheds light on the (systematic component of) heterogeneity in returns to diversification. Using data on excess values for a subset of Compustat firms over three decades, we show that (i) the standard deviation of systematic heterogeneity attributable to diversification is roughly twice as large as the mean discount, (ii) the standard deviation of these systematic effects amounts to 18-35% of the market value of a firm, and (iii) roughly one-third of firms display a systematic diversification premium. Framed differently, our estimates imply that the mean diversification discount (μ_D) accounts for only 9-19% of

[48] In the data, Coke refocused and its excess value improved whereas Walmart diversified and its excess value dropped.

total systematic variation in excess values due to diversification, while other firm-specific systematic factors related to diversification (μ_i) account for the remaining 81-91%. The results are robust to the inclusion of standard observable proxies that are employed in the literature, to outliers, and to controlling for the endogeneity of diversification choices.

Our approach separates systematic heterogeneity attributable to the diversification choice from other generalized sources of heterogeneity (for eg., firm level capabilities or resources unrelated to the diversification choice itself) and from somewhat persistent but ultimately transitory factors (for eg., strategic drivers that are imitable even if not in the very short-run). Each of these is an important source of persistent performance differences in the data. As a result, the misspecification from employing standard estimation approaches that do not separate these different sources can be large—for example, standard "firm fixed effect" approaches to estimating systematic heterogeneity dramatically overstate the actual magnitude of systematic heterogeneity in our data. Furthermore, we show that the role of diversification choices in systematic heterogeneity is comparable in magnitude to systematic heterogeneity from all other sources combined. This underscores the importance of diversification decisions—but for reasons stemming from systematic performance heterogeneity rather than the traditional logic related to a mean discount or premium.

The results have both important descriptive and prescriptive content. For one, shifting the focus of the diversification debate "from means to variances"—and, similarly, away from observable or easily measurable proxies that lend themselves to large-sample research toward understanding the root drivers of such differences even if they don't—is likely to be a promising area of inquiry. Perhaps as important, doing so would also be more consistent with theories of strategic advantage that are ultimately rooted in understanding such differences.[49] Beyond this, generalizing the approach offered here to understanding

[49] In a recent survey, Stein (2003) argues that focusing attention on heterogeneity in the diversification discount rather than on the mean discount is an approach more consistent with existing theory on corporate investment behavior. He notes that "after all, taken as a whole, the theoretical work does not lead to a clear-cut prediction that diversification is on average good or bad. Rather, the theory has more bite in the cross-section, pointing to the specific circumstances under which internal capital markets are most likely to destroy value. Thus, the diversification discount may indeed be a useful measure, but perhaps one should

the importance of systematic heterogeneity in explaining the value from other firm-level strategic choices—R&D, offshoring, alliances, and relocation—is likely to be fruitful in other areas of empirical strategy research.

References

[1] Adner, Ron, and Peter Zemsky, 2006, Diversification and Performance: Linking Relatedness, Market Structure and the Decision to Diversify, working paper.

[2] Anand, Bharat, and Tarun Khanna, 2000, Do Firms Learn to Create Value? The Case of Alliances, *Strategic Management Journal*, 21, 295-315.

[3] Barney, Jay, 1991, Firm Resources and Sustained Competitive Advantage, *Journal of Management*, 17(1), 99-120.

[4] Berger, Philip G., and Rebecca Hann, 2003, The Impact of SFAS No. 131 on Information and Monitoring, *Journal of Accounting Research*, 41(2), 163-223.

[5] Berger, Philip G., and Eli Ofek, 1995, Diversification's Effect on Firm Value, *Journal of Financial Economics,* 37, 39-65.

[6] Bernardo, Antonio, and Bhagwan Chowdhry, 2002, Resources, Real Options, and Corporate Strategy, *Journal of Financial Economics*, 63, 211-234.

[7] Bernstein, Jeffrey, and Ishaq Nadiri, 1988, Interindustry R&D Spillovers, Rates of Return, and Production in High-Tech Industries, *American Economic Review*, 78(2), 429-434.

[8] Bowman, Edward, and Constance Helfat, 2001, Does Corporate Strategy Matter? *Strategic Management Journal*, 22, 1-23.

[9] Brickley, James, and Leonard Van Drunen, 1990, Internal Corporate Restructuring: An Empirical Analysis, *Journal of Accounting and Economics*, 12, 251-280.

[10] Campa, José Manuel, and Simi Kedia, 2002, Explaining the diversification discount, *Journal of Finance,* 57, 1731-1762.

[11] Cassiman, Bruno, and Reinhilde Veugelers, 2006, In Search of Complementarity in the Innovation Strategy: Internal R&D and External Knowledge Acquisition, *Management Science*, 52(1), 68-82.

[12] Chan, Su Han, George Gau, and Ko Wang, 1995, Stock Market Reaction to Capital Investment Decisions: Evidence from Business Relocations, *Journal of Financial and Quantitative Analysis*, 30, 81-100.

pay less attention to its mean value, and more to its cross-sectional variation."

[13] Chan, Louis, Josef Lakonishok, and Theodore Sougiannis, 2001, The Stock Market Valuation of Research and Development Expenditures, *Journal of Finance*, 56(6), 2431-2456.

[14] Chang, Sea-Jin, and Harbir Singh, 2000, Corporate and Industry Effects on Business Unit Competitive Position, *Strategic Management Journal*, 21, 739-752.

[15] Chevalier, Judith, 2004. What Do We Know About Cross-Subsidization? Evidence from Merging Firms. *Advances in Economic Analysis and Policy*, 4(1), 1-27.

[16] Cockburn, Iain, and Zvi Griliches, 1988, Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents, *American Economic Review*, 78(2), 419-423.

[17] Collis, David J., and Cynthia A. Montgomery, 2005, *Corporate Strategy: A Resource-Based Approach*. 2nd ed. New York: McGraw-Hill/Irwin.

[18] Dyer, Jeff, Prashant Kale, and Harbir Singh, 2001, How to Make Strategic Alliances Work: The Role of the Alliance Function, *MIT Sloan Management Review*, July 2001.

[19] Erickson, Gary, and Robert Jacobson, 1992, Gaining Comparative Advantage Through Discretionary Expenditures: The Returns to R&D and Advertising, *Management Science*, 38(9), 1264-1279.

[20] Gertner, Robert, Comments in "Research Roundtable Discussion: The Diversification Discount." FEN-Educator Series, 2003.

[21] Gomes, Joao, and Dmitry Livdan, 2004, Optimal Diversification: Reconciling Theory and Evidence, *Journal of Finance*, 59(2), 507-535.

[22] Hall, Bronwyn, 1993, The Stock Market's Valuation of R&D Investment During the 1980's, *American Economic Review*, 83(2), 259-264.

[23] Kotha, Suresh, and Anil Nair, 1995, Strategy and Environment as Determinants of Performance: Evidence from the Japanese Machine Tool Industry, *Strategic Management Journal*, 16, 497-518.

[24] Lamont, Owen A., and Christopher Polk, 2001, Cash flows vs. returns, *Journal of Finance*, 56, 1693-1721.

[25] Lang, Larry H.P., and René M. Stulz, 1994, Tobin's q, Corporate Diversification, and Firm Performance, *Journal of Political Economy*, 102(6), 1248-1280.

[26] Levinthal, Daniel, and Brian Wu, 2010, Opportunity Costs and Non-Scale Free Capabilities: Profit Maximization, Corporate Scope, and Profit Margins, *Strategic Management Journal*, 31(7), 780-801.

[27] Lippman, Steven, and Richard Rumelt, 1982, Uncertain Imitability: an Analysis of Interfirm Differences in Efficiency under Competition, *Bell Journal of Economics*, 13, 418-438.

[28] McGahan, Anita, and Michael Porter, 1997, How much does Industry Matter, Really? *Strategic Management Journal*, 18, 15-30.

[29] McGahan, Anita, and Michael Porter, 1999, The Persistence of Shocks to Profitability, *Review of Economics and Statistics*, 81(1), 143–153.

[30] McGahan, Anita, and Michael Porter, 2002, What do we Know about Variance in Accounting Profitability? *Management Science*, 48(7), 834-851.

[31] McGahan, Anita, and Michael Porter, 2003, The Emergence and Sustainability of Abnormal Profits, *Strategic Organization*, 1, 79-108.

[32] McGahan, Anita, and Michael Porter, 2005, Comment on 'Industry, Corporate and Business-Segment Effects and Business Performance: A Non-Parametric Approach' by Ruefli and Wiggins, *Strategic Management Journal*, 26, 873-880.

[33] Meyer, Margaret, Paul Milgrom, and John Roberts, 1992, Organizational Prospects, Influence Costs, and Ownership Changes, *Journal of Economics and Management Strategy*, 1(1), 9-35.

[34] Morck, Randall, Andrei Shleifer, and Robert Vishny, 1990, Do Managerial Objectives Drive Bad Acquisitions? *Journal of Finance*, 45(1), 31-48.

[35] Peteraf, Margaret, 1993, The Cornerstones of Competitive Advantage: A Resource-Based View, *Strategic Management Journal*, 14(3), 179–191.

[36] Porter, Michael, 1996, What is Strategy, *Harvard Business Review*, Nov/Dec 1996.

[37] Prezas, Alexandros, Karen Simonyan, and Gopala Vasudevan, 2010, Does Offshoring Create Value for Shareholders? *Review of Financial Economics*, 19(4), 179-191.

[38] Rajan, Raghuram, Henri Servaes, and Luigi Zingales, 2000, The cost of diversity: The diversification discount and inefficient investment, *Journal of Finance*, 55, 35-80.

[39] Rivkin, Jan, 2000, Imitation of Complex Strategies, *Management Science*, 46(6), 824-844.

[40] Rumelt, Richard, 1991, How much does Industry Matter? *Strategic Management Journal*, 12, 167-185.

[41] Ryall, Michael, 2009, Causal Ambiguity, Complexity, and Capability-Based Advantage, *Management Science*, 55(3), 389-403.

[42] Sampson, Rachelle, 2004, Experience Effects and Collaborative Returns in R&D Alliances, *Strategic Management Journal*, 26(11), 1009-1031.

[43] Sampson, Rachelle, 2007, R&D Alliances and Firm Performance: The Impact of Technological Diversity and Alliance Organization on Innovation, *The Academy of Management Journal*, 50(2), 364-386.

[44] Schmalensee, Richard, 1985, Do markets differ much? *American Economic Review*, 75, 341-351.

[45] Spanos, Yiannis, George Zaralis, and Spyros Lioukas, 2004, Strategy and Industry Effects on Profitability: Evidence from Greece, *Strategic Management Journal*, 25, 139-165.

[46] Stein, Jeremy C., 2003, Agency Information, and Corporate Investment, Constantinides, George, Milton Harris, and René Stulz, Eds., *Handbook of the Economics of Finance*, Amsterdam: North Holland.

[47] Stulz, René, 1990, Managerial Discretion and Optimal Financing Policies, *Journal of Financial Economics*, 26, 3-27.

[48] Villalonga, Belén, 2004a, Does Diversification Cause the "Diversification Discount"?, *Financial Management,* 33, 5-27.

[49] Villalonga, Belén, 2004b, Diversification Discount or Premium? New Evidence from the Business Information Tracking Series, *Journal of Finance*, 59, 475–502.

[50] Wernerfelt, Birger, 1984, The Resource-Based View of the Firm, *Strategic Management Journal*, 5(2), 171–180.

Table 1. Descriptive statistics

	asset multipliers		sales multipliers		industry-adjusted *q*-s	
	1978-1996	**1998-2008**	**1978-1996**	**1998-2008**	**1978-1996**	**1998-2008**
mean diversification discount	-0.111	-0.101	-0.124	-0.115	-0.137	-0.119
median diversification discount	-0.105	-0.096	-0.130	-0.135	-0.102	-0.082
SD(excess values):						
diversified firms	0.476	0.645	0.608	0.853	0.459	0.647
single-segment firms	0.577	0.739	0.726	0.968	0.604	0.805
% diversified firm-years that have higher excess values than the median single-segment firm	38%	42%	39%	40%	37%	43%

The sample selection criteria are described in section 2.
The mean (median) diversification discount is computed as the difference in mean (median) excess values between single-segment and diversified firms.
The excess value measures based on asset multipliers and sales multipliers follow Berger and Ofek (1995). We compute them as log(market value/imputed value). Market value is the market value of common equity plus book value of debt and preferred equity. The imputed value of a firm is the sum of the imputed values of its segments, where a segment's imputed value is equal to the segment assets (sales) multiplied by the median ratio of market value to assets (sales) for the single-segment firms in the corresponding industry-year. The excess value measure based on industry-adjusted *q*-s follows Lang and Stulz (1994). We compute it as the percentage difference between a firm's actual *q* and its imputed *q*. The imputed *q* for a firm is the asset-weighted average of the imputed *q*-s of its segments, where a segment's imputed *q* is computed as the average *q* for the single-segment firms in the corresponding industry-year, and *q* is measured as the ratio of market to book value.

Table 2. Fixed-effects estimates of firm-specific diversification effects and firm effects

The estimation model is

$$Y_{i,t} = g_t + a_i + m_i D_{i,t} + v_{i,t},$$

where $Y_{i,t}$ is excess value for firm i in year t, $D_{i,t}$ is a dummy variable equal to one if the firm is diversified (multi-segment), g_t is the year effect, a_i is the firm effect, m_i is the firm-specific diversification effect, and $v_{i,t}$ is the error term. The model is estimated using OLS, and a_i and m_i are firm-specific free parameters in estimation.

	asset multipliers		sales multipliers		industry-adjusted *q*-s	
	1978-1996	**1998-2008**	**1978-1996**	**1978-1996**	**1998-2008**	**1978-1996**
SD(diversification effects m_i)	0.464	0.626	0.570	0.754	0.462	0.761
SD(firm effects a_i)	0.465	0.579	0.610	0.808	0.485	0.614
cor(m_i, a_i)	-0.580	-0.575	-0.530	-0.469	-0.604	-0.592

The sample selection criteria are described in section 2.
The excess value measures based on asset multipliers and sales multipliers follow Berger and Ofek (1995). We compute them as log(market value/imputed value). Market value is the market value of common equity plus book value of debt and preferred equity. The imputed value of a firm is the sum of the imputed values of its segments, where a segment's imputed value is equal to the segment assets (sales) multiplied by the median ratio of market value to assets (sales) for the single-segment firms in the corresponding industry-year. The excess value measure based on industry-adjusted *q*-s follows Lang and Stulz (1994). We compute it as the percentage difference between a firm's actual *q* and its imputed *q*. The imputed *q* for a firm is the asset-weighted average of the imputed *q*-s of its segments, where a segment's imputed *q* is computed as the average *q* for the single-segment firms in the corresponding industry-year, and *q* is measured as the ratio of market to book value.

Table 3. Summary of the baseline estimates

The estimation model is

$$Y_{i,t} = \lambda_t+\alpha_i+(\mu_D+\mu_i)D_{i,t}+u_{i,t}$$

where $Y_{i,t}$ is excess value for firm i in year t, $D_{i,t}$ is a dummy variable equal to one if the firm is diversified (multi-segment), λ_t is the year effect, α_i is the systematic firm effect, μ_D is the mean diversification discount or premium, μ_i is the systematic firm-specific diversification effect, and $u_{i,t}$ is the error term that represents possibly persistent but not permanent shocks modeled as AR(K). The model is estimated using maximum likelihood, and the optimal number of lags K is chosen based on Bayes Information Criterion (BIC). The firm effects α_i and the firm-specific diversification effects μ_i are random effects, i.e., we estimate the parameters of their joint distribution.

	asset multipliers		**sales multipliers**		**industry-adjusted *q*-s**	
	1978-1996	**1998-2008**	**1978-1996**	**1998-2008**	**1978-1996**	**1998-2008**
μ_D	-0.13	-0.10	-0.16	-0.14	-0.15	-0.11
SD(μ_i)	0.23	0.18	0.34	0.29	0.34	0.30
SD(α_i)	0.29	0.35	0.43	0.57	0.31	0.33
cor(μ_i, α_i)	-0.42	0.04	-0.47	0.23	-0.62	0.21
Pr($\mu_D+\mu_i$ >0)	28%	29%	32%	31%	33%	35%
Pr($\alpha_i+\mu_D+\mu_i$ >0)	32%	40%	35%	42%	29%	41%
log likelihood	-23596.1	-24557	-31747.5	-29489.1	-23391.9	-26950.8
# firm-years	52803	29730	52803	29730	52803	29730
# firms	7052	5412	7052	5412	7052	5412
best # lags	2	2	2	2	2	2

The sample selection criteria are described in section 2.

The excess value measures based on asset multipliers and sales multipliers follow Berger and Ofek (1995). We compute them as log(market value/imputed value). Market value is the market value of common equity plus book value of debt and preferred equity. The imputed value of a firm is the sum of the imputed values of its segments, where a segment's imputed value is equal to the segment assets (sales) multiplied by the median ratio of market value to assets (sales) for the single-segment firms in the corresponding industry-year. The excess value measure based on industry-adjusted *q*-s follows Lang and Stulz (1994). We compute it as the percentage difference between a firm's actual *q* and its imputed *q*. The imputed *q* for a firm is the asset-weighted average of the imputed *q*-s of its segments, where a segment's imputed *q* is computed as the average *q* for the single-segment firms in the corresponding industry-year, and *q* is measured as the ratio of market to book value.

Table 4. Robustness check: summary of estimates after adding control variables from Berger and Ofek (1995) and related diversification

The estimation model is

$$Y_{i,t} = \lambda_t + \beta X_{i,t} + \alpha_i + (\mu_0 + \mu_1 RelDiv_{i,t} + \mu_i) D_{i,t} + u_{i,t},$$

where $Y_{i,t}$ is excess value for firm i in year t, $X_{i,t}$ are the control variables, $D_{i,t}$ is a dummy variable equal to one if the firm is diversified (multi-segment), $RelDiv_{i,t}$ is the related diversification variable, λ_t is the year effect, α_i is the systematic firm effect, μ_0 and μ_1 are parameters capturing the mean value consequences of diversification, μ_i is the systematic firm-specific diversification effect, and $u_{i,t}$ is the error term that represents possibly persistent but not permanent shocks modeled as AR(K).
The control variables $X_{i,t}$ are log assets, EBIT/sales and CAPEX/sales, following Berger and Ofek (1995). The related diversification variable $RelDiv_{i,t}$ is computed as follows. First, for each firm-year, we merge segments that have the same 2-digit SIC code. After that, for each firm-year, we compute $RelDiv_{i,t}$ as the Herfindahl index of its sales across these merged segments.
The model is estimated using maximum likelihood, and the optimal number of lags K is chosen based on Bayes Information Criterion (BIC). The firm effects α_i and the firm-specific diversification effects μ_i are random effects, i.e., we estimate the parameters of their joint distribution.

	asset multipliers		**sales multipliers**		**industry-adjusted *q*-s**	
	1978-1996	**1998-2008**	**1978-1996**	**1978-1996**	**1998-2008**	**1978-1996**
$\mu_D = \mu_0 + \mu_1 RelDiv$ *	-0.11	-0.12	-0.22	-0.24	-0.12	-0.10
$SD(\mu_i)$	0.22	0.18	0.33	0.30	0.33	0.30
$SD(\alpha_i)$	0.29	0.35	0.41	0.53	0.32	0.34
$cor(\mu_i, \alpha_i)$	-0.42	0.06	-0.53	0.21	-0.58	0.24
$Pr(\mu_D + \mu_i > 0)$ *	31%	25%	26%	21%	36%	36%
$Pr(\alpha_i + \mu_D + \mu_i > 0)$ *	34%	38%	28%	36%	35%	42%
log likelihood	-22574.4	-24478.9	-30644.6	-28931	-22513.7	-26901.2
# firm-years	52803	29730	52803	29730	52803	29730
# firms	7052	5412	7052	5412	7052	5412
best # lags	2	2	2	2	2	2

* computed at the average level of related diversification.
The sample selection criteria are described in section 2.
The excess value measures based on asset multipliers and sales multipliers follow Berger and Ofek (1995). We compute them as log(market value/imputed value). Market value is the market value of common equity plus book value of debt and preferred equity. The imputed value of a firm is the sum of the imputed values of its segments, where a segment's imputed value is equal to the segment assets (sales) multiplied by the median ratio of market value to assets (sales) for the single-segment firms in the corresponding industry-year. The excess value measure based on industry-adjusted *q*-s follows Lang and Stulz (1994). We compute it as the percentage difference between a firm's actual *q* and its imputed *q*. The imputed *q* for a firm is the asset-weighted average of the imputed *q*-s of its segments, where a segment's imputed *q* is computed as the average *q* for the single-segment firms in the corresponding industry-year, and *q* is measured as the ratio of market to book value.

Table 5. Robustness check: summary of estimates after controlling for strategic proxy variables and related diversification

The estimation model is

$$Y_{i,t} = \lambda_t + \beta X_{i,t} + \alpha_i + (\mu_0 + \mu_1 RelDiv_{i,t} + \mu_i) D_{i,t} + u_{i,t},$$

where $Y_{i,t}$ is excess value for firm i in year t, $X_{i,t}$ are the strategic proxy variables, $D_{i,t}$ is a dummy variable equal to one if the firm is diversified (multi-segment), $RelDiv_{i,t}$ is the related diversification variable, λ_t is the year effect, α_i is the systematic firm effect, μ_0 and μ_1 are parameters capturing the mean value consequences of diversification, μ_i is the systematic firm-specific diversification effect, and $u_{i,t}$ is the error term that represents possibly persistent but not permanent shocks modeled as AR(K).
The strategic proxy variables $X_{i,t}$ are: R&D intensity and advertising intensity as proxies for differentiation; market share and log total assets as proxies for economies of scale/scope; and assets/sales and CAPEX/sales as measures of "asset parsimony" (Kotha and Nair 1995). We industry-adjust all strategic variables by computing them as deviations of the original variables from the imputed industry medians for each year. We also include the same industry-adjusted variables squared. The related diversification variable $RelDiv_{i,t}$ is computed as follows. First, for each firm-year, we merge segments that have the same 2-digit SIC code. After that, for each firm-year, we compute $RelDiv_{i,t}$ as the Herfindahl index of its sales across these merged segments.
The model is estimated using maximum likelihood, and the optimal number of lags K is chosen based on Bayes Information Criterion (BIC). The firm effects α_i and the firm-specific diversification effects μ_i are random effects, i.e., we estimate the parameters of their joint distribution.

	asset multipliers		**sales multipliers**		**industry-adjusted *q*-s**	
	1978-1996	**1998-2008**	**1978-1996**	**1978-1996**	**1998-2008**	**1978-1996**
$\mu_D = \mu_0 + \mu_1 RelDiv$ *	-0.13	-0.12	-0.17	-0.16	-0.14	-0.11
$SD(\mu_i)$	0.22	0.18	0.27	0.24	0.33	0.29
$SD(\alpha_i)$	0.29	0.34	0.33	0.43	0.33	0.35
$cor(\mu_i, \alpha_i)$	-0.42	0.05	-0.44	0.06	-0.60	0.23
$Pr(\mu_D+\mu_i>0)$ *	28%	26%	27%	26%	34%	35%
$Pr(\alpha_i+\mu_D+\mu_i>0)$ *	32%	39%	30%	38%	32%	41%
log likelihood	-23410.3	-24512.9	-27786.4	-27256.6	-23111.9	-26887.8
# firm-years	52803	29730	52803	29730	52803	29730
# firms	7052	5412	7052	5412	7052	5412
best # lags	2	2	2	2	2	2

* computed at the average level of related diversification.
The sample selection criteria are described in section 2.
The excess value measures based on asset multipliers and sales multipliers follow Berger and Ofek (1995). We compute them as log(market value/imputed value). Market value is the market value of common equity plus book value of debt and preferred equity. The imputed value of a firm is the sum of the imputed values of its segments, where a segment's imputed value is equal to the segment assets (sales) multiplied by the median ratio of market value to assets (sales) for the single-segment firms in the corresponding industry-year. The excess value measure based on industry-adjusted *q*-s follows Lang and Stulz (1994). We compute it as the percentage difference between a firm's actual *q* and its imputed *q*. The imputed *q* for a firm is the asset-weighted average of the imputed *q*-s of its segments, where a segment's imputed *q* is computed as the average *q* for the single-segment firms in the corresponding industry-year, and *q* is measured as the ratio of market to book value.

Table 6. Robustness check: summary of estimates for switchers

The estimation model is

$$Y_{i,t} = \lambda_t + \alpha_i + (\mu_D + \mu_i) D_{i,t} + u_{i,t}$$

where $Y_{i,t}$ is excess value for firm i in year t, $D_{i,t}$ is a dummy variable equal to one if the firm is diversified (multi-segment), λ_t is the year effect, α_i is the systematic firm effect, μ_D is the mean diversification discount or premium, μ_i is the systematic firm-specific diversification effect, and $u_{i,t}$ is the error term that represents possibly persistent but not permanent shocks modeled as AR(K). The model is estimated using maximum likelihood, and the optimal number of lags K is chosen based on Bayes Information Criterion (BIC). The firm effects α_i and the firm-specific diversification effects μ_i are random effects, i.e., we estimate the parameters of their joint distribution.

The sample consists of firms that diversified or refocused during the sample period.

	asset multipliers		**sales multipliers**		**industry-adjusted *q*-s**	
	1978-1996	**1998-2008**	**1978-1996**	**1978-1996**	**1998-2008**	**1978-1996**
μ_D	-0.08	-0.07	-0.11	-0.08	-0.08	-0.07
SD(μ_i)	0.21	0.16	0.31	0.25	0.33	0.28
SD(α_i)	0.32	0.24	0.44	0.38	0.35	0.30
cor(μ_i, α_i)	-0.45	0.12	-0.45	0.01	-0.63	-0.38
Pr(μ_D+μ_i >0)	35%	34%	36%	37%	40%	41%
Pr(α_i+μ_D+μ_i >0)	39%	41%	40%	43%	39%	42%
log likelihood	-5126.4	-3796.88	-7026.08	-4679.46	-4383.34	-4205.45
# firm-years	11607	4816	11607	4816	11607	4816
# firms	1101	781	1101	781	1101	781
best # lags	2	2	2	2	2	2

The sample selection criteria are described in section 2, and applied to the subsample of firms that diversified or refocused during the sample period.

The excess value measures based on asset multipliers and sales multipliers follow Berger and Ofek (1995). We compute them as log(market value/imputed value). Market value is the market value of common equity plus book value of debt and preferred equity. The imputed value of a firm is the sum of the imputed values of its segments, where a segment's imputed value is equal to the segment assets (sales) multiplied by the median ratio of market value to assets (sales) for the single-segment firms in the corresponding industry-year. The excess value measure based on industry-adjusted *q*-s follows Lang and Stulz (1994). We compute it as the percentage difference between a firm's actual *q* and its imputed *q*. The imputed *q* for a firm is the asset-weighted average of the imputed *q*-s of its segments, where a segment's imputed *q* is computed as the average *q* for the single-segment firms in the corresponding industry-year, and *q* is measured as the ratio of market to book value.

Table 7. Robustness check: summary of estimates after the correction for selection following Campa and Kedia (2002)

The estimation model in the second stage is

$$Y_{i,t} = \lambda_t+\beta X_{i,t}+\alpha_i+(\mu_D+\mu_i)D_{i,t}+\gamma IMR_{i,t}+u_{i,t}$$

where $Y_{i,t}$ is excess value for firm *i* in year *t*, $D_{i,t}$ is a dummy variable equal to one if the firm is diversified (multi-segment), $X_{i,t}$ are the control variables, $IMR_{i,t}$ is the inverse Mill's ratio from stage 1, λ_t is the year effect, α_i is the systematic firm effect, μ_D is the mean diversification discount or premium, μ_i is the systematic firm-specific diversification effect, and $u_{i,t}$ is the error term that represents possibly persistent but not permanent shocks modeled as AR(K).

The control variables $X_{i,t}$ are log total assets for years t, t-1, t-2, EBIT/sales for years t, t-1, t-2, CAPEX/sales for years t, t-1, t-2, leverage and log total assets squared (the control variables follow Campa and Kedia 2002). In the first stage, we run a probit regression of the diversification dummy on the control variables $X_{i,t}$, year dummies, and instruments for diversification: the proportion of other diversified firms within the 2-digit industry of firm *i* in year *t*, the proportion of other firms' sales within the 2-digit industry of firm *i* that is due to other diversified firms, S&P index dummy, a major stock exchange dummy and a foreign incorporation dummy (the instruments follow Campa and Kedia 2002).

The second-stage model is estimated using maximum likelihood, and the optimal number of lags K is chosen based on Bayes Information Criterion (BIC). The firm effects α_i and the firm-specific diversification effects μ_i are random effects, i.e., we estimate the parameters of their joint distribution.

	asset multipliers		**sales multipliers**		**industry-adjusted *q*-s**	
	1978-1996	**1998-2008**	**1978-1996**	**1978-1996**	**1998-2008**	**1978-1996**
μ_D	0.02	0.21	-0.01	0.18	0.01	0.14
SD(μ_i)	0.20	0.19	0.29	0.29	0.30	0.35
SD(α_i)	0.28	0.35	0.38	0.52	0.29	0.33
cor(μ_i , α_i)	-0.44	0.13	-0.55	-0.20	-0.59	0.30
γ	-0.062	-0.189	-0.085	-0.210	-0.051	-0.143
Pr($\mu_D+\mu_i>0$)	54%	86%	48%	73%	51%	66%
Pr($\alpha_i+\mu_D+\mu_i>0$)	53%	69%	49%	63%	51%	60%
log likelihood	-15443.4	-20419	-20797.3	-23588.7	-14006	-21372.3
# firm-years	45195	26777	45195	26777	45195	26777
# firms	5991	4907	5991	4907	5991	4907
best # lags	2	2	2	2	2	2

The sample selection criteria are described in section 2. The number of observations is smaller than in previous tables due to missing values for lagged control variables.

The excess value measures based on asset multipliers and sales multipliers follow Berger and Ofek (1995). We compute them as log(market value/imputed value). Market value is the market value of common equity plus book value of debt and preferred equity. The imputed value of a firm is the sum of the imputed values of its segments, where a segment's imputed value is equal to the segment assets (sales) multiplied by the median ratio of market value to assets (sales) for the single-segment firms in the corresponding industry-year. The excess value measure based on industry-adjusted *q*-s follows Lang and Stulz (1994). We compute it as the percentage difference between a firm's actual *q* and its imputed *q*. The imputed *q* for a firm is the asset-weighted average of the imputed *q*-s of its segments, where a segment's imputed *q* is computed as the average *q* for the single-segment firms in the corresponding industry-year, and *q* is measured as the ratio of market to book value.

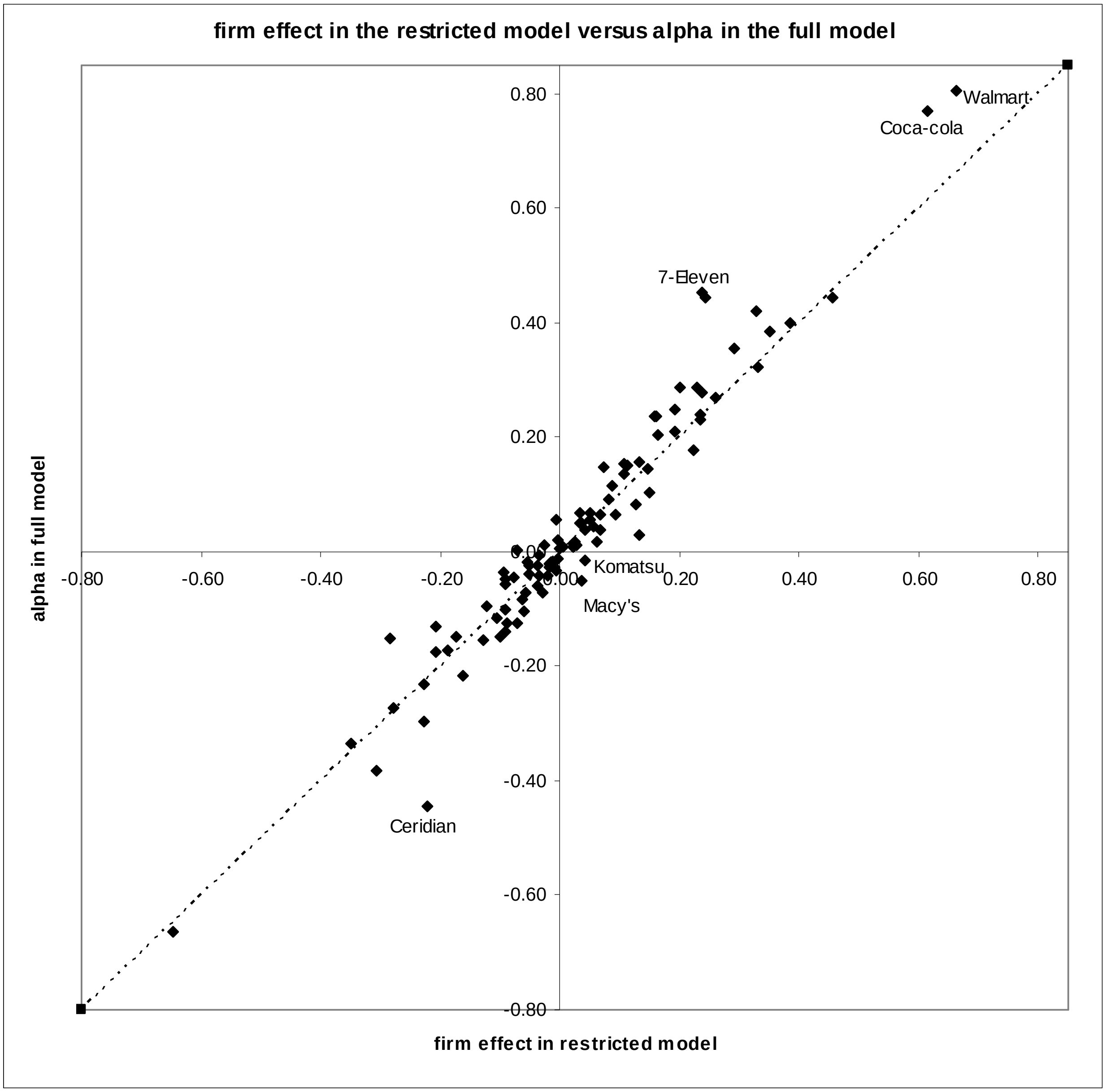

Figure 1. Firm effects versus "generic" systematic heterogeneity α_i.

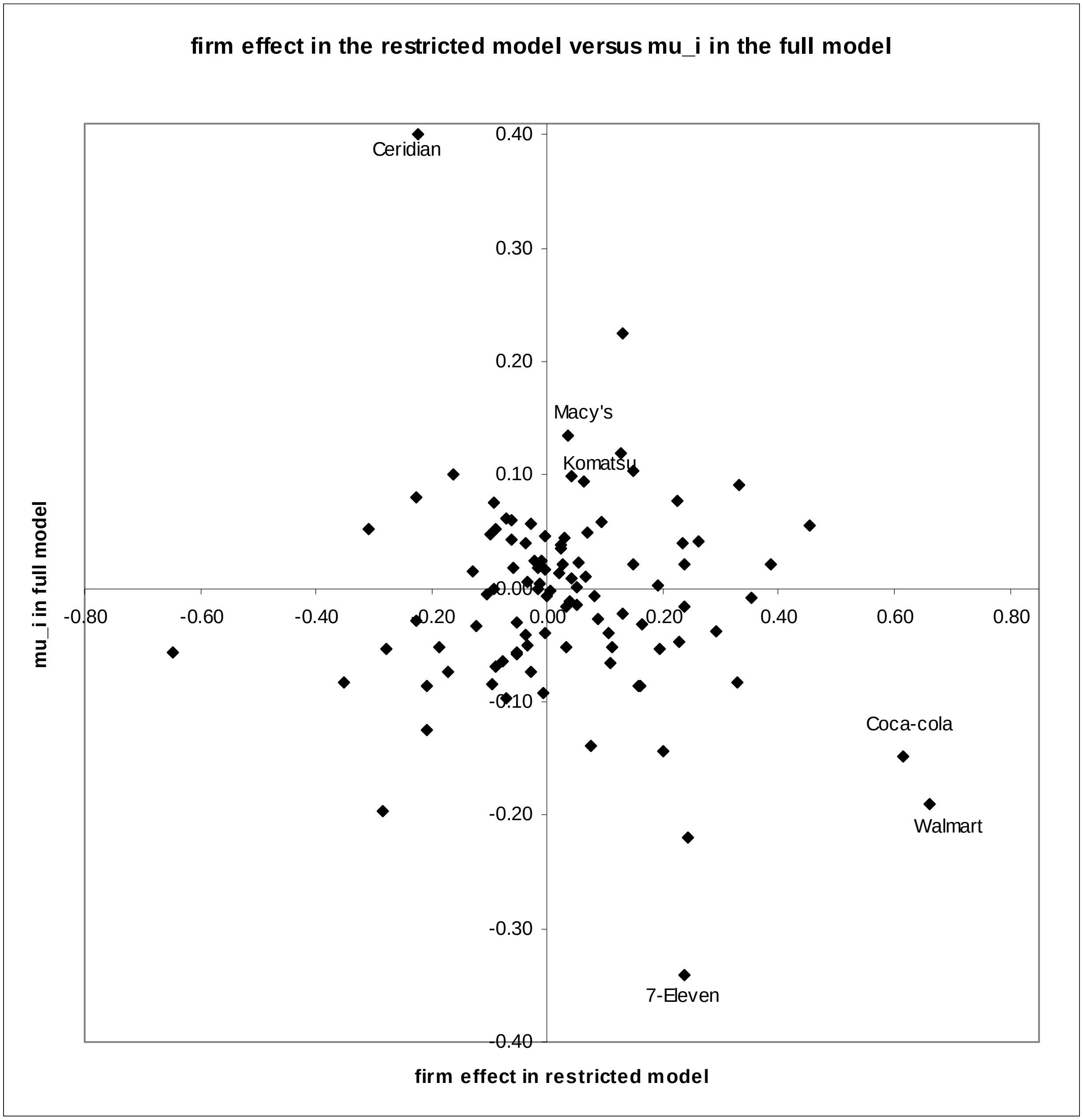

Figure 2. Firm effects versus diversification-related systematic heterogeneity μ_i.