THE JOURNAL OF FINANCE • VOL. LVI, NO. 5 • OCT. 2001

The Diversification Discount: Cash Flows Versus Returns

OWEN A. LAMONT and CHRISTOPHER POLK*

ABSTRACT

Diversified firms have different values from comparable portfolios of single-segment firms. These value differences must be due to differences in either future cash flows or future returns. Expected security returns on diversified firms vary systematically with relative value. Discount firms have significantly higher subsequent returns than premium firms. Slightly more than half of the cross-sectional variation in excess values is due to variation in expected future cash flows, with the remainder due to variation in expected future returns and to covariation between cash flows and returns.

THE VALUE OF ANY ASSET DEPENDS mechanically on the asset's future cash flows and future returns. This fact is, by definition, true both for actual future cash flows and returns and for expected future cash flows and returns. High future cash flows imply high value today, and high future returns imply low value today. Thus, to explain the value of an asset, one needs to examine expected future cash flows and expected future returns.

In recent years, the average diversified firm has been worth less than a portfolio of comparable single-segment firms (Lang and Stulz (1994), Berger and Ofek (1995)). A large literature attempts to explain this fact by exploring ways that diversification might affect cash flows. This literature hypothesizes that diversification itself causes the diversified firm to generate different cash flows than it would if separated into single-segment firms. Potential explanations include incompetent or irrational managers, competent but self-interested managers, wasteful spending in general, and wasteful investment in poorly performing divisions in particular. See, for example, Morck, Shleifer, and Vishny (1990), Comment and Jarrell (1995), Servaes (1996), Denis, Denis, and Sarin (1997), Lamont (1997), Scharfstein (1998), Rajan, Servaes, and Zingales (2000), and Scharfstein and Stein (2000).

* Owen A. Lamont is at the Graduate School of Business, University of Chicago. Christopher Polk is at the Kellogg Graduate School of Management, Northwestern University. We thank Eugene F. Fama, Charles M. Jones, Darius Palia, Raghuram Rajan, Henri Servaes, Jeremy Stein, René Stultz, Tuomo Vuolteenaho, Luigi Zingales, two anonymous reviewers, and seminar participants at Chicago, Duke, Kellogg, the NBER, Virginia Tech, and Wharton for helpful comments. We thank Eugene F. Fama for providing data. Lamont gratefully acknowledges support from the Alfred P. Sloan Foundation, the Center for Research in Securities Prices at the University of Chicago Graduate School of Business, and the National Science Foundation.

A second explanation is that diversification does not affect value, but rather merely reflects patterns in what types of firms tend to agglomerate together into diversified firms. If firms generating lower cash flow tend to cluster together into conglomerates, then the fact that the average conglomerate is worth less than a comparable portfolio of single-segment firms does not necessarily imply value destruction (see, for example, Chevalier (1999) and Maksimovic and Phillips (1999)).

Both these explanations implicitly assume that returns are the same for diversified firms and single-segment firms. If returns are the same, then two portfolios can only have different values if the future cash flows are different. We examine a third explanation for the diversification discount: Diversified firms have expected future asset returns that are different from the returns of single-segment firms. Different securities can have different expected returns for many reasons; explanations include risk, mispricing, taxes, and liquidity.

To study why the *average level* of diversified firm values is low, one type of evidence used in the previous literature is the *cross-sectional distribution* of diversified firm values. For example, Berger and Ofek (1995) run regressions with firm value on the left-hand side, and investment and cross-subsidization variables on the right-hand side. They then use these cross-sectional regressions to make inferences about the average level of discount. Our paper is also about the cross-sectional distribution of diversified firm value. Although our paper provides no direct evidence explaining the average level of the discount, one can infer that the same mechanism that explains the cross section might also explain the average level.

Specifically, we perform a variance decomposition for the cross-sectional distribution of diversified firm value, and quantify the relative importance of cash flow and returns. Our approach is based on an identity relating value to future cash flows and returns. Other things being equal, a diversified firm with a high expected return (relative to single-segment firms) will have a low value and thus a discount. A diversified firm with relatively low expected return will have a premium. We test whether variation in excess values is explainable using variation in expected returns. We examine the difference in subsequent returns on diversified firms and on single-segment firms. We find that excess values forecast future returns in the required way. Firms with discounts have higher subsequent returns than firms with premia. The diversification discount puzzle is, at least in part, an expected return phenomenon as well as an expected cash flow phenomenon.

This paper is organized as follows. Section I shows the basic identity we use. Section II describes the sample and shows summary statistics. Section III examines monthly portfolio returns and shows basic results on return predictability. Section IV briefly examines three explanations for the different returns: risk, liquidity, and mispricing. Section V examines present value relations using annual data on firm returns, and shows what fraction of cross-sectional variation in excess values is due to different returns and what fraction is due to different cash flows. Section VI summarizes and presents conclusions.

I. The Identity

To decompose the diversification discount into differences in cash flows and differences in returns, we use the fact that the price of any asset is the sum of its discounted future dividends, based on the definition of returns:

$$R_t = \frac{D_{t+1} + P_{t+1} - P_t}{P_t}, \tag{1}$$

where D_t is dividend paid out during period t and P_t is price at the end of period t. "Dividends" includes all cash flows paid to the security holders (including interest payments made to bondholders). Equation (1) defines returns for any portfolio or asset, including a firm's equity, a portfolio of a firm's equity and debt, or a portfolio of many firms' securities. Iterating forward and imposing a terminal condition on the growth of stock prices in the infinite future,

$$P_t = \sum_{j=1}^{\infty} \frac{D_{t+j}}{\prod_{k=1}^{j}(1 + R_{t+k})}. \tag{2}$$

Equation (2) holds for realized returns and realized cash flows. Again, one can also take expectations of both sides and say that the value of any asset depends on expectations of the interaction of future cash flows and future returns. We describe later how to disentangle expected returns and expected cash flows.

We define the excess value on a diversified firm as the log ratio of the value of a diversified firm and the value of a portfolio of single-segment firms, $\ln(P_t/\bar{P}_t)$, where the bar indicates single segment. Using equation (2), excess value is:

$$\ln\left(\frac{P_t}{\bar{P}_t}\right) = \ln\left(\sum_{j=1}^{\infty} \frac{D_{t+j}}{\prod_{k=1}^{j}(1 + R_{t+j})}\right) - \ln\left(\sum_{j=1}^{\infty} \frac{\bar{D}_{t+j}}{\prod_{k=1}^{j}(1 + \bar{R}_{t+j})}\right). \tag{3}$$

In calculating excess value, we use a portfolio of single-segment companies and normalize this portfolio to have the same level of either sales or book assets as the diversified firm. Thus the price and dividends on the single-segment portfolio have been multiplied by (for example) the ratio of the diversified firm's current sales to single-segment portfolio current sales. A negative excess value is a diversification discount and a positive excess value is a diversification premium. Equation (3) shows that, mechanically, the diversification discount depends on future cash flows and future returns. One can also take the expectation of both sides and say that the excess value depends on expected future cash flows and expected future returns.

II. The Sample

The sample consists of firms reporting segment data in the COMPUSTAT Current and Research database, 1979 through 1997. For each segment, firms report sales, book assets, and other variables. In addition, COMPUSTAT assigns each segment a four-digit SIC code based on the line of business description of the segment. We define a diversified firm as a firm with at least two segments, and a single-segment firm as a firm with only one segment. Following Berger and Ofek (1995), we discard firm-years with segments in the financial services industry, total firm sales of less than \$20 million, or discrepancies between segment and firm data. See the Appendix for more details.

We use two measures of value. The first is Q, the market–book ratio of the firm, calculated as the ratio of the market value of the firm divided by the total book assets of the firm. The second is M, the market–sales ratio of the firm, calculated as the ratio of the market value of the firm to total sales of the firm.

For each value measure, we calculate the ratio for a comparable portfolio of single-segment firms. Again following Berger and Ofek (1995), for each segment of a diversified firm, we find a group of matching single-segment firms with similar SIC codes. We match either by two-, three-, or four-digit SIC code, using the highest precision that meets the requirement of having at least five single-segment firms in a given year. We then calculate the value measure for each segment, using either the weighted average or the median. For weights, we use either the book value of assets for Q or sales for M. We then form a value measure for the entire diversified firm, dropping every diversified firm that does not have comparable value measures for each of its segments.

For a given diversified firm with n different segments, the comparable ratios for a given year are

$$\bar{Q} = \sum_{j=1}^{n} wa_j \left(\frac{1}{\sum_{i=1}^{N_j} A_i} \sum_{i=1}^{N_j} A_i Q_i \right), \tag{4}$$

$$\bar{Q}_{MEDIAN} = \sum_{j=1}^{n} wa_j (median\{Q_1, Q_2, \ldots, Q_{N_j}\}), \tag{5}$$

$$\overline{M} = \sum_{j=1}^{n} ws_j \left(\frac{1}{\sum_{i=1}^{N_j} S_i} \sum_{i=1}^{N_j} S_i M_i \right), \tag{6}$$

$$\overline{M}_{MEDIAN} = \sum_{j=1}^{n} ws_j (median\{M_1, M_2, \ldots, M_{N_j}\}) \tag{7}$$

Table I

Value Ratios and Leverage Ratios for Diversified Firms, 1979 to 1997

Lower case letters indicate natural logarithm. Q is the market–book ratio. M is the market–sales ratio. D is the debt ratio, defined as the book value of the debt divided by the book value of the debt plus the market value of the equity. Comparable portfolio variables are

$$\bar{Q} = \sum_{j=1}^{n} wa_j \left[\left(\sum_{i=1}^{N_j} A_i \right)^{-1} \sum_{i=1}^{N_j} A_i Q_i \right], \qquad \bar{Q}_{MEDIAN} = \sum_{j=1}^{n} wa_j (median\{Q_1, Q_2, \ldots, Q_{N_j}\}),$$

$$\text{and} \qquad \bar{D}_{ASSET\ WEIGHT} = \sum_{j=1}^{n} wa_j \left[\left(\sum_{i=1}^{N_j} A_i \right)^{-1} \sum_{i=1}^{N_j} A_i D_i \right],$$

where wa_j is the fraction of the firm's assets that are in segment j; Q_i and A_i are the q ratio and book assets of single-segment company i; and segment j's industry has N_j single-segment firms. The comparable portfolio variables for M are defined similarly using sales weights. The sample consists of all diversified firms for which at least five matches could be found for every segment. "Sales Weight" indicates weighting by total sales of the diversified firm. There are 14,962 annual observations.

Variable	Mean	Median	Standard Deviation	Minimum	Maximum	Mean, Sales Weight	Fraction Positive
Q	1.35	1.15	0.72	0.35	16.27	1.46	1.00
$\bar{Q}$	1.56	1.44	0.56	0.64	8.05	1.61	1.00
$\bar{Q}_{MEDIAN}$	1.34	1.24	0.42	0.62	6.15	1.37	1.00
M	1.12	0.75	1.44	0.02	57.20	1.11	1.00
$\bar{M}$	1.28	1.06	0.87	0.11	12.98	1.28	1.00
$\bar{M}_{MEDIAN}$	1.09	0.87	0.79	0.08	12.74	1.08	1.00
$q - \bar{q}$	−0.18	−0.19	0.39	−2.02	2.17	−0.12	0.28
Premium firms only	0.26	0.18	0.26	0.00	2.17	0.22	1.00
Discount firms only	−0.36	−0.31	0.27	−2.02	0.00	−0.30	0.00
$q - \bar{q}_{MEDIAN}$	−0.05	−0.07	0.36	−1.94	2.31	0.02	0.40
$m - \bar{m}$	−0.30	−0.29	0.63	−3.77	2.97	−0.23	0.30
Premium firms only	0.40	0.29	0.38	0.00	2.97	0.29	1.00
Discount firms only	−0.60	−0.51	0.45	−3.77	0.00	−0.48	0.00
$m - \bar{m}_{MEDIAN}$	−0.13	−0.13	0.61	−3.78	3.16	−0.05	0.40
$D - \bar{D}_{ASSET\ WEIGHT}$	0.05	0.02	0.23	−0.71	0.83	0.00	0.53
$\bar{D} - \bar{D}_{SALES\ WEIGHT}$	0.06	0.02	0.22	−0.70	0.83	0.01	0.55

where wa_j is the fraction of the firm's assets that are in segment j; ws_j is the fraction of the firm's sales that are in segment j; Q_i, M_i, A_i, and S_i are the q ratio, market–sales ratio, book assets, and sales of single-segment company i; and segment j's industry has N_j single-segment firms.

Table I shows summary statistics for value ratios, excess values, leverage, and returns. Lower case letters indicate natural logs. Table I's sample contains 14,962 annual observations for 2,390 different diversified firms in the

19-year period of 1979 through 1997. The average number of segments per firm is 2.8. Each segment of a diversified firm is matched with an average of 11 single-segment firms.

Table I contains several different ways of calculating excess values. One way is to compare the levels, for example, to subtract mean Q and $\bar{Q}$ to obtain a mean excess value of -0.21. A second way is to calculate the excess values by taking the natural logarithm of the ratio of the Qs or Ms, as in Berger and Ofek (1995). We focus on log ratios because they are an important component of the present value calculations performed in Section V.

Measured with log ratios, the average and median excess values are negative, and range from -5 to -30 percent, similar to previous research.[1] Excess value is positive for about a third of the sample (the fraction ranges from 28 percent to 40 percent across the different measures). We show median excess values in Table I only for comparison to previous research. In this paper, we necessarily concentrate on average excess values, since we need to calculate returns on the portfolio of single-segment firms. Table I also shows the average excess values calculated by weighting each observation by the diversified firms' sales. Size-weighting tends to raise the average excess value. Since most of the literature on the diversification discount uses equal weighting (as any OLS regression does), we focus on equal weighting in this paper.

Table I gives a sense of the cross-sectional distribution of excess values by showing summary statistics for premium firms and discount firms. Looking at average $q - \bar{q}$, premium firms have substantial premiums of 26 percent, while discount firms have substantial discounts of 36 percent. The table also shows a wide distribution of excess values within premium firms and discount firms.

Table I shows that diversified firms have higher debt ratios than single-segment firms. The debt ratio, D_t, is defined as the end of year t ratio of the book value of debt to the book value of debt plus the market value of equity (where debt includes preferred stock). In calculating leverage ratios for portfolios of single-segment firms, we again weight either by book assets or by sales.

To calculate returns on the diversified firm as a whole, in principle one needs both equity and debt returns. We obtain equity returns for each firm from the Center for Research in Securities Prices (CRSP). Unfortunately, debt returns are not available for most firms. Just using equity returns would be unwise, since leverage is somewhat higher for diversified firms than for single-segment firms (as shown in Table I). It also turns out that discount firms have higher debt ratios than premium firms. Thus leverage is a potentially important confounding factor.

[1] In interpreting the values it is important to note that logarithms are concave functions. Since firm-level variables are more volatile than industry-level variables, average log ratios tend to be negative. For example, mean Q is above mean $\bar{Q}_{MEDIAN}$, but mean $q - \bar{q}_{MEDIAN}$ is negative.

We therefore approximate debt returns for each firm using returns on the Lehman Brothers Corporate Bond Index. For year t, we define total firm returns as

$$R_t^{FIRM\,J} = D_{t-1}^{J} R_t^{AGGREGATE\,BOND\,RETURNS} + (1 - D_{t-1}^{J}) R_t^{FIRM\,J\,EQUITY}. \tag{8}$$

This method of calculating total returns introduces three biases into our analysis. The first two biases are in favor of the null hypothesis, while the third seems ambiguous. First, discount firms are more levered and thus probably have riskier debt with higher expected return. Consistent with this idea, Hecht (1999) finds a small negative relationship between firm market-to-book ratios and subsequent bond returns. By using aggregate bond returns, we are understating total returns on discount firms and overstating total returns on premium firms. Second, discount firms may have debt that has deteriorated in value and has market value below book value, so that the calculated leverage ratios overstate actual leverage (relative to premium firms). Since average equity returns are higher than average debt returns, we are again understating discount firm total returns and overstating premium firm total returns. Because we intend to test whether total returns on discount firms are higher than total returns on premium firms, we are conservatively measuring returns in a way that is biased against our hypothesis. Third, if the diversified firms and matching single segment firms have debt that is less (more) risky than aggregate, the fact that discount firms are more levered will bias their returns up (down). The effect of this third bias is unclear.

III. Monthly Portfolio Returns

We now test the basic hypothesis of this paper, that excess values are related to expected security returns on diversified firms. We test whether realized future returns on discount firms are significantly higher than realized future returns on premium firms, and discuss evidence that the patterns in realized returns reflect patterns in expected returns. In this section, we test the hypothesis using simple portfolio formation rules; later, in Section V, we test the hypothesis using regression methods.

Our portfolio formation rules follow Fama and French (1993) and incorporate realistic timing constraints. The basic algorithm is that each July of year t, one sorts firms into portfolios based on information in December of year $t-1$. One examines returns on this portfolio from July of year t until June of year $t+1$. We use this timing convention to ensure that the sorting information is certainly in the information set of investors.

Table II shows average monthly returns on diversified firms. Panel A reports total raw returns (i.e., using no information about the returns on single-segment firms). Panels B and C look at excess returns, defined as diversified

Table II

Time-series Average Monthly Returns for Diversified Firms, 1980 to 1998

Average monthly returns in percent, 1980:7–1998:12. Excess returns are firm returns minus industry benchmark returns. Total returns are a linear combination, using firm leverage, of returns on firm equity and returns on aggregate corporate debt. Panel A has an average of 838 diversified firms per month, and Panels B to C have an average of 714 diversified firms per month. In each row, we report 10 portfolio strategies. Columns (1) and (10) take positions in all diversified firms. Columns (2) through (5) only take positions in premium firms, as determined by excess values measured using either firms' market–book ratios (q) or market–sales ratios (m). Similarly, columns (6) through (9) only take positions in discount firms, as determined by either q or m. "Top half" means that the firm's excess value is above the median excess value for that group and year. Standard errors are in parentheses.

		Premium Firms				Discount Firms				Difference
	All Firms	All	Top Half	Bottom Half	Top minus Bottom	All	Top Half	Bottom Half	Top minus Bottom	All Premium minus All Discount
	(1)	(2)	(3)	(4)	(5): (4) − (3)	(6)	(7)	(8)	(9): (8) − (7)	(10): (6) − (2)
Panel A: R: Total raw returns										
Using q,	1.19	0.97	0.88	1.05	0.17	1.27	1.20	1.34	0.14	0.30
asset weights	(0.23)	(0.26)	(0.30)	(0.23)	(0.11)	(0.22)	(0.22)	(0.22)	(0.05)	(0.08)
Using m,	1.19	0.99	0.89	1.09	0.20	1.27	1.17	1.37	0.20	0.28
sales weights	(0.23)	(0.24)	(0.26)	(0.23)	(0.08)	(0.22)	(0.22)	(0.23)	(0.07)	(0.06)
Panel B: $R - \bar{R}$: Total excess returns										
Using q,	−0.04	−0.25	−0.36	−0.14	0.22	0.04	0.01	0.07	0.06	0.29
asset weights	(0.07)	(0.09)	(0.12)	(0.09)	(0.10)	(0.08)	(0.08)	(0.11)	(0.08)	(0.11)
Using m,	−0.07	−0.25	−0.33	−0.17	0.16	0.00	−0.08	0.09	0.17	0.26
sales weights	(0.07)	(0.08)	(0.10)	(0.09)	(0.08)	(0.08)	(0.08)	(0.10)	(0.07)	(0.08)
Panel C: $R - \bar{R}$: Equity excess returns										
Using q,	0.02	−0.34	−0.45	−0.24	0.21	0.15	0.07	0.24	0.17	0.50
asset weights	(0.06)	(0.09)	(0.12)	(0.09)	(0.11)	(0.08)	(0.07)	(0.11)	(0.11)	(0.11)
Using m,	−0.02	−0.32	−0.42	−0.23	0.20	0.10	−0.07	0.27	0.34	0.43
sales weights	(0.06)	(0.08)	(0.11)	(0.09)	(0.12)	(0.08)	(0.07)	(0.11)	(0.10)	(0.11)

firm returns minus returns on the portfolio of comparable single-segment firms. We focus on excess returns because we are trying to explain excess value using returns on diversified firms compared to single-segment firms.

To calculate excess returns, we go short the portfolio of comparable single-segment firms for each diversified firm, weighting the firms in the same manner used in constructing excess value. In calculating firm returns, we use year $t - 1$ data on sales, assets, debt ratios, and SIC codes. The result is a zero-cost portfolio of excess (or industry-adjusted) returns called $R - \bar{R}$, diversified firm returns minus returns on the industry benchmark.[2] Our sample consists of diversified firms (approximately 714 per month) for which at least five matching single-segment firms could be found for each segment. There are two versions of $R - \bar{R}$, one based on asset weights and one based on sales weights, which correspond to the two ways of defining excess value. One could describe $R - \bar{R}$ as the return from a "pairs" trading strategy of buying diversified firms, and shorting similar single-segment firms (see, for example, Gatev, Goetzmann, and Rouwenhorst, 1999).

Panel B shows total excess returns. We start by discussing column (1) in Panel B, which shows excess returns for all diversified firms. To explain the average diversification discount using future returns, one would need a positive and significant excess return in column (1) of Panel B. Instead, Table II shows that diversified firms have excess returns that are about zero during the sample period. In other words, diversified firms have returns that are the same as the portfolio of comparable single-segment firms. Thus, column (1) of Panel B fails to explain the average diversification discount. One explanation for this result is that, over short time periods such as our sample, realized returns are a noisy measure of expected returns (a point made forcefully by Elton (1999)). For example, if diversified firms had positive expected excess returns, but unexpected bad news occurred in the sample period, the negative unexpected returns could swamp the positive expected returns.

Rather than looking at excess returns on all diversified firms, a more powerful test is to compare returns on premium firms and discount firms. This test is more powerful because the return differential on these two groups is unaffected by events that affect both groups equally. In the context of diversified firms, Lang and Stulz (1994) argue that the ex post performance of diversified firms is a potentially misleading measure of ex ante valuation because of (p. 1253) "unexpected technological and regulatory changes." If these unexpected shocks are common to all diversified firms, they do not affect the difference between premium firm returns and discount firm returns.

We now turn to whether discount firms have returns that are higher than premium firms. Each July of year t, we sort firms into portfolios based on

[2] We do not require that either the diversified firm or the single-segment firm is present for the entire period. If a firm exits from the CRSP database, we drop it from the portfolio using the delisting return. Using the Shumway (1997) adjustment does not substantially change the results in Table II. For example, the average differential returns in Column (10) never change more than one basis point, and remain positive and significant.

their excess values as of December of year $t - 1$. We construct two portfolios, a portfolio that buys premium firms and a portfolio that buys discount firms. Using raw returns, column (10) of Panel A shows that discount firms have total returns that are 30 basis points per month higher than premium firms sorting on q, and 28 basis points higher sorting on m. These differences are statistically significant.

Column (10) of Panel B shows the basic results of this paper: premium firms have significantly lower excess returns than discount firms. Sorting by q, column (2) shows that premium firms have returns that are 25 basis points per month lower than a comparable portfolio of single-segment firms, while column (6) shows that discount firms have returns that are four basis points higher than single-segment firms. The mean excess return of premium firms is significantly different from zero; the mean excess return of discount firms is not. More importantly, column (10) shows that the difference of 29 basis points in the excess returns is significantly different from zero. Sorting by m, the difference is 26 basis points per month. Panel C shows, as expected, the difference between premium returns and discount returns increases when using equity returns (rather than total firm returns).

Looking more closely at the cross-sectional distribution, columns (3) to (5) and (6) to (9) look within premium and discount firms. Each year, we rank all premium firms and discount firms by excess value and split each group in half. For example, the bottom half of discount firms, column (8), consists of deep discount firms that have low excess values (below the median excess value for discount firms in that year). This split has the property that, for example, returns on all discount firms in column (6) equals the average of the returns in columns (7) and (8), so that one can see the components of the various returns.

We then test whether, within groups, the top half has higher returns than the bottom half. Column (5) shows that, within premium firms, high premium firms have returns that are 16 to 22 basis points higher than low premium firms, significantly higher in five out of six cases. Column (9) shows that within discount firms, high discount firms have returns that are 6 to 34 basis points lower than low discount firms, significantly lower in four out of six cases. More dramatically, one can compare deep discount firms in column (8) to high premium firms in column (3). The last row of Table II shows, for example, that deep discount firms have significantly positive excess returns, high premium firms have significantly negative excess returns, and the difference is 69 basis points a month, or more than eight percent per year.

In summary, Table II shows that excess values forecast excess returns, so that at least part of the diversification discount phenomenon can be explained by future returns. The difference between premium and discount firm returns is statistically and economically (at three to six percent per year) significant. Looking within premium and discount firms, the pattern remains consistent. Higher excess values today mean lower subsequent returns.

The returns in Table II are equal weighted in the sense that each diversified firm in each month has the same weight in calculating the month's returns. They are partially size weighted in the sense that the single-segment firms are always weighted either by sales or book assets when forming the zero cost portfolio $R - \bar{R}$. An alternative method of calculating monthly returns would be to size weight each individual $R - \bar{R}$ by the diversified firms' sales. When we size weight the returns in Table II (not shown), we find that the differential return is lower and insignificant.[3] Since our goal is to explain how much of the cross-sectional variation in excess value is due to return differences, we believe equal weighting is appropriate in our context. We are trying to relate the diversification discount to predictable variation in returns, and to understand the extensive literature on the diversification discount in light of this relation. Since the previous literature on the diversification discount uses equal weighting, we do the same.

The present value equation (equation (2)) is true for realized returns, by definition. A more interesting question for financial economists is whether the cross-sectional return patterns reflect expected returns, where "expected" means predictable in advance by a rational agent. Here we present evidence that suggests that the variation in returns documented in Table II was predictable ex ante, and did not merely reflect ex post realizations that happened to appear during the sample period.

We examine the pattern of differential returns earned over time. If the differential returns were concentrated in one specific time period, it would suggest that the differential returns just happened to occur during our sample, and could have been a surprise to investors. If the differential returns were consistently positive year after year, it would suggest that expected returns, not surprises, explain the positive differential. A specific story about surprises is that during the takeover wave of the late 1980s, many diversified firms divested unrelated subsidiaries, experienced bust-up takeovers, went private, or took other value-enhancing actions (e.g., Comment and Jarrell (1995), Berger and Ofek (1996)). Berger and Ofek (1996) find that takeovers and LBOs for diversified firms peaked in 1988, and fell sharply thereafter, suggesting that the differential return should fall after 1988.

Table III displays evidence on the time series of annual differential excess returns from column (10) of Table II, Panel B. For each of the 18 years, we report total excess returns for the 12-month period ending in June. Sorting by q, differential returns are positive in 14 years (sorting by m, not shown, differential returns are positive in 15 years). The pattern appears just as strong in the 1990s as in the 1980s. The time pattern suggests that higher returns on discount firms were not just lucky draws that surprised investors.

More general evidence from other research also supports the idea that these return patterns are not random. The pattern in returns in diversified firms is an example of the more general "value effect" in security returns:

[3] For the six differential returns reported in column (10) of Table II, sales-weighting produces a mean differential of between −0.06 and 0.18. The differential is negative two out of six times. All six estimates are insignificant.

Table III

Excess Value and Annual Differential Returns on Diversified Firms, 1979 to 1998

"Number of observations" corresponds to the number of firms in year $t-1$ (not necessarily in the returns from year t to year $t+1$). Statistics on excess value, $q-\bar{q}$, are from the end of year $t-1$. $(R-\bar{R})_{\mathrm{DISCOUNT}}-(R-\bar{R})_{\mathrm{PREMIUM}}$ is annual percent returns on the total excess return differential strategy in column (10) of Table II, Panel B.

			Mean $q-\bar{q}$				
Year $t-1$	Number of Obs.	Fraction Positive	All Firms	Premium Firms	Discount Firms	Spread	$(R-\bar{R})_{\mathrm{DISCOUNT}}-(R-\bar{R})_{\mathrm{PREMIUM}}$ July Year t–June Year $t+1$
	(1)	(2)	(3)	(4)	(5)	(6): (4) − (5)	(7)
79	862	0.29	−0.13	0.23	−0.28	0.51	−9.21
80	1,031	0.31	−0.13	0.31	−0.33	0.64	18.62
81	960	0.27	−0.19	0.25	−0.34	0.59	−4.52
82	939	0.27	−0.22	0.26	−0.39	0.65	15.62
83	895	0.24	−0.22	0.25	−0.38	0.63	3.53
84	898	0.26	−0.18	0.23	−0.32	0.55	0.49
85	870	0.27	−0.20	0.24	−0.37	0.61	6.59
86	790	0.29	−0.20	0.22	−0.36	0.58	5.74
87	786	0.28	−0.18	0.24	−0.35	0.59	3.89
88	737	0.30	−0.14	0.24	−0.31	0.55	−2.73
89	683	0.30	−0.17	0.26	−0.35	0.61	−3.96
90	686	0.29	−0.19	0.27	−0.37	0.64	5.56
91	695	0.27	−0.24	0.33	−0.45	0.78	5.16
92	703	0.29	−0.20	0.30	−0.40	0.70	9.51
93	713	0.30	−0.19	0.29	−0.39	0.68	2.14
94	736	0.30	−0.18	0.26	−0.36	0.62	4.30
95	731	0.28	−0.20	0.30	−0.40	0.70	3.47
96	705	0.30	−0.20	0.28	−0.41	0.69	7.53
97	542	0.34	−0.17	0.28	−0.41	0.51	

subsequent returns are negatively correlated with value levels. For example, looking prior to our sample period, Davis, Fama, and French (2000) show a consistent value effect in U.S. equities going back to the 1920s.

Table III also provides additional information on the distribution of excess values over time. It shows the total number of firms per year, the percent with premiums, the average excess value, and the average excess value for premium and discount firms. There is no apparent trend in these numbers, except for the number of diversified firms, which shrinks over time (falling from a peak of 1,031 in 1980 to a low of 542 in 1997). The discount for all diversified firms is fairly constant, ranging from 13 to 24 percent.

One revealing pattern in Table III is the relation between excess values and subsequent differential returns. The spread between excess values in premium and discount firms varies over time. If this variation reflects changes in expected returns, then the excess value differential in year $t - 1$ should be positively correlated with excess return differential in year t. That is, in years when premium firms have higher prices relative to discount firms, subsequent returns on premium firms should be lower relative to discount firms. In fact, this pattern is evident in the 18 annual observations of Table III. The correlation between excess value differentials and subsequent excess return differentials is positive and significant at 0.50. Thus a profitable time to buy discount firms and short premium firms is when discount firms are especially cheap and premium firms are especially expensive.[4]

Having documented that there is substantial variation in expected returns across diversified firms, we next turn to explaining the sources of this return predictability.

IV. Risk, Liquidity, and Mispricing

In judging *whether* expected returns drive the diversification discount, it is not necessary to establish *why* expected returns on diversified firms and single-segment firms differ. This question is of independent interest, however. One explanation for our results is the value effect: Stocks with high scaled prices have low subsequent returns. At least since Ball (1978), financial economists have argued that scaled price should contain information about future returns. Researchers have documented this effect in various contexts ranging from closed-end funds to international equities. Our contribution is to document a case of this value effect in order that the valuation of diversified firms is not misinterpreted as arising solely from differences in cash flows.

Explaining the value effect is beyond the scope of this paper, but in this section we take a first pass at three explanations for the predictability of diversified firm returns. First, we examine multifactor explanations based

[4] Similarly, Cohen, Polk, and Vuolteenaho (2000) find that the spread in the book-to-market ratio of a value strategy predicts its subsequent return.

Table IV

Three-factor Regressions on Equity Excess Returns, 1980 to 1997

$$(R - \bar{R})_{\text{DISCOUNT}} - (R - \bar{R})_{\text{PREMIUM}} = a + bRMRF + sSMB + hHML$$

The dependent variable is the difference between equity excess returns on discount firms and premium firms, from column (10) of Table II, Panel C. The independent variables are from Fama and French (1993) and include *RMRF*, the market return minus the risk-free return; *SMB*, the size factor; and *HML*, the market-to-book factor. The mean repeats information from Table II, which has sample 1980:7–1998:12. The regression results reflect the sample 1980:7–1997:12. Standard errors are in parentheses.

	Mean	a	b	s	h	R^2
$(R - \bar{R})_{\text{DISCOUNT}} - (R - \bar{R})_{\text{PREMIUM}}$	0.50	0.41	−0.01	0.17	0.37	0.33
sorted on $q - \bar{q}$	(0.11)	(0.10)	(0.03)	(0.04)	(0.04)	
$(R - \bar{R})_{\text{DISCOUNT}} - (R - \bar{R})_{\text{PREMIUM}}$	0.43	0.32	0.02	0.19	0.38	0.36
sorted on $m - \bar{m}$	(0.11)	(0.09)	(0.02)	(0.04)	(0.04)	

on risk. Second, we examine whether differences in liquidity explain excess values. Third, we examine the possibility of mispricing related to liquidity costs of trading.

A. Multifactor Risk Explanations

Table IV examines in more detail the differential returns earned by portfolio strategies that buy discount equities (and short the comparable single-segment equities) and short premium equities (and buy the comparable single-segment equities). It tests whether the Fama–French (1993) three-factor model can explain these differential returns. We use equity returns, not total firm returns, since that is what the three-factor model is designed to explain.

The first column shows the mean return on this strategy (the same number reported in column (10) of Table II, Panel C). The regression results in the rest of the table show that while the differential return loads positively on the value factor, HML, the three-factor model does not explain these equal-weighted differential returns very well. Sorting by q, the three-factor model explains only 9 out of the 50 basis points of the return differential. Sorting by m, the three-factor model describes only 11 of the 43 basis points of the return differential. Thus the patterns in diversified firm stock returns do not merely reflect loadings on the value factor.

B. Liquidity

Capozza and Seguin (1999) reach similar conclusions to this paper. Based on a study of the real estate industry, they conclude that the diversification discount is not due to differences in cash flows, and so must be due to differences in required return. They find suggestive evidence that differences in liquidity (measured by equity trading volume) are positively related to differences in excess value. The idea is that investors demand higher expected returns to compensate for illiquidity.

An implication of Capozza and Seguin (1999), in line with Amihud and Mendelson (1986), is that one should observe high excess returns when diversified firms are less liquid than their comparable portfolio of single-segment firms, and low excess returns when diversified firms are more liquid. We test this implication using our sample of diversified firm excess returns. Following Capozza and Seguin (1999), we use the dollar volume of the firm's equity to measure liquidity. Since we look at equity volume, we examine equity returns. Again, we form portfolios in July based on information on liquidity and excess value as of the previous December.

Due to the differences between specialist and dealer markets, an additional Nasdaq trade increases reported trading volume by two shares while an additional NYSE/AMEX trade increases reported trading volume by only one share. To create a common measure of volume across exchanges, we use total dollar volume for NYSE/AMEX firms and one-half total dollar volume for Nasdaq firms. Using this adjustment, for each diversified firm, we calculate both its adjusted dollar volume and the adjusted dollar volume of the portfolio of matching single-segment firms, where the portfolio weighting again uses either assets or sales.[5]

The left-hand side of Table V shows results for liquidity and returns. It tests whether differences in returns are monotonically related to differences in liquidity. We sort both diversified firms and their particular matching single-segment portfolios into three liquidity groups, and calculate average monthly returns for the resulting nine configurations. For example, the upper left-hand corner of the table shows average excess returns on diversified firms for which both the firm and its matching portfolio are low liquidity positions. According to the hypothesis that return differences are a function of liquidity differences, excess returns should be decreasing as one moves northeast in this half of the table.

Table V shows that the predicted relation between excess returns and liquidity differences is basically present. As predicted by the hypothesis, the lower left corner has higher average excess returns than the upper right corner. Though the differences between those two corners are economically large for both sorts, those differences are not statistically significant in either case. When sorting by q, the difference is 26 basis points with a t-statistic of 0.76. The m sorts generate a difference of 29 basis points with a t-statistic of 0.87.

[5] Single-segment firms are slightly more likely to trade on Nasdaq. The fraction that trade on Nasdaq is 30 percent for diversified firms and 32 percent for the matching portfolio using asset weights (33 percent using sales weights). This two percent difference is the same for both premium and discount firms measured using q (and the three percent difference using sales weights is also the same for both premium and discount firms measured using m).

Table V

Dollar Volume and Monthly Returns on Diversified Firms, 1980 to 1998

Dollar volume is the average of all CRSP available monthly dollar trading volume of a firm from January $t-1$ to December $t-1$. For Nasdaq firms, we divide the average CRSP dollar volume in half due to Nasdaq reporting standards. We then calculate the (asset or sales-weighted) matched dollar volume for each diversified firm's single-segment matching portfolio. Each year we sort all diversified firms on the diversified firm's dollar volume into three portfolios. We independently sort all diversified firms on the diversified firm's matched portfolio dollar volume. From the intersection of these two sorts we form nine portfolios. We then calculate the equal-weighted excess return over the period July t to June $t+1$ on these portfolios as well as the difference between the excess returns on the premium and discount subsets within each portfolio. We report below the time-series average return on these portfolios. The nine portfolios have an average of 74 stocks over the 19-year period with approximately 70 percent of the diversified firms within each of the nine portfolios being discount firms. The sample period is 1980:7–1998:12. Nasdaq firms are in the sample from 1984:7 to 1998:12 as a full year of volume information for those firms becomes available on CRSP in 1983. Standard errors are in parentheses.

	$R-\bar{R}$			$(R-\bar{R})_{\text{DISCOUNT}} - (R-\bar{R})_{\text{PREMIUM}}$		
	Diversified Firm Volume					
	Low	Medium	High	Low	Medium	High
Based on q and asset weights						
Single segment firm volume						
Low	0.17	−0.12	−0.18	0.46	0.44	0.12
	(0.17)	(0.12)	(0.16)	(0.31)	(0.23)	(0.20)
Medium	0.34	0.14	−0.05	0.71	0.50	0.22
	(0.18)	(0.11)	(0.12)	(0.41)	(0.27)	(0.18)
High	0.08	−0.05	−0.12	0.72	0.71	0.18
	(0.23)	(0.14)	(0.09)	(0.39)	(0.24)	(0.14)
Based on m and sales weights						
Single segment firm volume						
Low	0.12	−0.19	−0.19	0.54	0.45	0.32
	(0.17)	(0.13)	(0.17)	(0.29)	(0.20)	(0.20)
Medium	0.27	0.03	−0.09	0.53	0.46	0.26
	(0.19)	(0.11)	(0.12)	(0.39)	(0.22)	(0.16)
High	0.10	−0.01	−0.12	−0.01	0.40	0.27
	(0.22)	(0.14)	(0.09)	(0.48)	(0.26)	(0.15)

C. Mispricing

An alternative hypothesis that explains our results is mispricing. A version of this hypothesis also has implications for liquidity and returns. If mispricing is more severe when it is difficult to arbitrage the mispriced assets, measures of arbitrage costs should be related to the predictability of returns. Pontiff (1996) shows that closed-end fund excess values are farther from zero when trading costs are higher (see also Shleifer and Vishny (1997)). Here, we assume that liquidity is negatively related to arbitrage costs and test whether portfolios of illiquid securities have more predictable returns.

In contrast to the hypothesis that liquidity is monotonically related to returns, the costly arbitrage view implies that as the illiquidity of either the diversified stocks or the comparable single-segment firms rise, the predictability of returns should rise. Returns should be most predictable (based on the level of excess value) when illiquidity makes it most difficult to take advantage of the mispricing.

The right half of Table V shows tests of the costly arbitrage hypothesis. It shows average excess return differentials between discount firms and premium firms. For example, the upper left-hand corner shows the difference between excess returns on discount firms and premium firms, where both sets of diversified firms and their matching portfolios have low liquidity. According to the hypothesis that returns become more predictable as illiquidity increases, differential excess returns should be decreasing as one moves southeast in this part of the table.

Table V shows that the predicted relation between differential excess returns and liquidity differences is weak at best. The hypothesis implies that the upper left corner should have higher average differential excess returns than the lower right corner, which is true for both sorting methods. However, the differences are statistically insignificant (the q sorts produce a difference of 28 basis points with a t-statistic of 0.84 while sorting by m generates a difference of 0.27 with a t-statistic of 0.89). Moreover, there is no obvious pattern of decreasing differential returns as one moves southeast across all nine portfolios.

In summary, we find no statistically convincing evidence linking liquidity-based explanations suggested by Capozza and Seguin (1999) and Amihud and Mendelson (1986) to our results. We also find no evidence supporting costly arbitrage explanations like those in Pontiff (1996). However, both investigations are certainly far less complete than previous work. We do find that differential returns are related to returns on the value factor of Fama and French (1993), but not well explained by their model. Thus we are unable to answer the question of why expected returns on diversified firms and single-segment firms differ; we are only able to document that they do differ.

V. Present Value Relations

In this section, we study the variance of excess values and use a dynamic model of returns and value ratios to decompose the cross-sectional variance into components due to cash flow and returns. The variance of excess values, $\mathrm{Var}(q_t - \bar{q}_t)$ or $\mathrm{Var}(m_t - \bar{m}_t)$, is the cross-sectional variance across all firm-years (which is shown in standard deviation form in Table I). This is the same object of interest in any regression with excess values as the dependent variable, as performed in many papers on the diversification discount.

The Campbell and Shiller (1988) log-linear approximation to the definition of return in equation (1) is

$$r_{t+1} \approx \rho p_{t+1} + (1-\rho) d_{t+1} - p_t + k, \tag{9}$$

where r is a continuously compounded return and lower case letters indicate natural logs. k is a constant coming from the approximation, which drops out below.

The parameter ρ is the inverse of one plus the ratio of dividends to market value, and is a discounting parameter that is close to one. In the context of this paper, dividends include interest payments. Using our assumptions about corporate debt returns, we calculated ρ and found it to be 0.96 for both diversified firms and for single-segment firms.[6] Thus we use 0.96 in our calculations.

Equation (9) holds either for diversified firms or for portfolios of single-segment firms. Subtracting the two,

$$r_{t+1} - \bar{r}_{t+1} = \rho(p_{t+1} - \bar{p}_{t+1}) + (1-\rho)(d_{t+1} - \bar{d}_{t+1}) - (p_t - \bar{p}_t). \tag{10}$$

In equation (10), one can scale the portfolio of single-segment firms so that it has the same level of sales or assets as the diversified firm, by multiplying prices and dividends by the ratio of the scaling variables. This renormalization has no effect on the left side of the equation, and allows one to use value ratios instead of actual prices in equation (10). It also means that dividends should be interpreted as the ratio of dividends to sales or to assets.

Iterating (10) forward and taking expected values of both sides, excess value is

$$p_t - \bar{p}_t = (1-\rho)E_t \sum_{j=0}^{\infty} \rho^j (d_{t+j+1} - \bar{d}_{t+j+1}) - E_t \sum_{j=0}^{\infty} \rho^j (r_{t+j+1} - \bar{r}_{t+j+1}). \tag{11}$$

This equation imposes the condition that the log dividend price ratio does not follow an explosive process. Introducing some notation, equation (11) can be rewritten as

$$p_t - \bar{p}_t = p_t^d - p_t^r. \tag{12}$$

Excess values consist of two parts. The first, p^d, is the sum of discounted future excess dividends (multiplied by $1-\rho$). The second, p^r, enters with a negative sign and is the sum of discounted future excess returns. Equation (11) is a completely atheoretical approximation to a dynamic accounting identity. It does not assume that financial markets are efficient or that market

[6] For each firm, we calculated the ratio of annual cash flow to end-of-year market value using dividend payments and interest payments. For interest payments, we used the firm's leverage ratio and the income component of the Lehman Brothers Corporate Bond Index. We found that average ρ was 0.956 for diversified firms and 0.962 for the comparable portfolio of single-segment firms (using either asset weighting or sales weighting). Campbell (1991) uses a monthly ρ of 0.9962 that translates into 0.955 annually for the aggregate stock market.

participants are rational. The terms "expected returns" and "expected dividends" refer to the rational expectation of returns and dividends, where the rational person is the econometrician not necessarily the investor.

A. Estimating the Dynamic Behavior of Discounts and Returns

Section III showed that excess values are related to subsequent returns using the standard, nonparametric, portfolio approach. In contrast, in this section we take a highly parametric approach that imposes homogeneity across all firms and years. We model the evolution of returns and value ratios using a vector autoregression (VAR). Let

$$\mathbf{x}'_{\mathbf{t}} = [r_t, \bar{r}_t, p_t, \bar{p}_t] \tag{13}$$

be the vector of returns and value ratios. We can represent the joint time-series behavior of returns and excess values using a first-order VAR:

$$\mathbf{x}_{\mathbf{t+1}} = \mathbf{c} + \mathbf{A}\mathbf{x}_{\mathbf{t}} + \mathbf{e}_{\mathbf{t+1}}, \tag{14}$$

where $\mathbf{A}$ is a 4×4 matrix of coefficients, $\mathbf{c}$ is a 4×1 vector of constants, and $\mathbf{e}$ is a 4×1 vector of error terms. Define $\mathbf{e_1}$ as the vector $[1\ \ 0\ \ 0\ \ 0]'$ and $\mathbf{e_2}$ as $[0\ \ 1\ \ 0\ \ 0]'$. After matrix algebra using equation (14), one can calculate the sum of discounted expected returns as

$$p_t^r = (\mathbf{e}'_1 - \mathbf{e}'_2)(I - \rho\mathbf{A})^{-1}\mathbf{A}\mathbf{x}_{\mathbf{t}} - (\mathbf{e}'_1 - \mathbf{e}'_2)(I - \rho\mathbf{A})^{-1}(\mathbf{I} - \rho\mathbf{I})^{-1}\mathbf{c}. \tag{15}$$

Using equation (12), p^d is then simply calculated as $p - \bar{p} + p^r$.

The second term in equation (15) is a constant term that is the same for all firms. Since our goal is to examine the variance of excess values across firms, the second term in equation (15) plays no role.

We estimate $\mathbf{A}$ using an annual VAR with log value ratios and continuously compounded returns. The VAR is estimated using four OLS regressions. The regressions require that the firm has annual returns and excess value ratios in both year $t + 1$ and year t (so that the firm must exist from the end of year $t - 1$ to the end of year $t + 1$). The data requirements cut the sample size substantially, compared to Table II.

Table VI shows VAR results. The first row, for example, shows coefficients from an OLS regression of annual diversified firm continuously compounded total returns on lagged returns and lagged value ratios, where the regression pools all firm-years. Again, each firm's total return is the weighted average of returns on the firm's equity and aggregate bond returns, using the firm's beginning-of-year debt ratio. The standard errors have been adjusted for correlation of the residuals within years, and for heteroskedasticity.[7]

[7] The robust standard errors allow for clustered sampling (dependence of observations within each year). See Rogers (1993).

Table VI

Dynamic Behavior of Annual Returns and Values Ratios for Diversified Firms and Matching Portfolios, 1981–1997

Vector autoregression results using pooled OLS estimation. The regression is $\mathbf{x_{t+1}} = \mathbf{c} + \mathbf{Ax_t} + \mathbf{e_{t+1}}$ where $\mathbf{x'_t} = [r_t, \bar{r}_t, p_t, \bar{p}_t]$ is the vector of continuously compounded returns and log value ratios. The sample consists of 8,698 diversified firm-years that have excess values and excess returns for two consecutive years. Using the estimated coefficients, we forecast future returns for each firm and calculate p^d and p^r. p^d is a function of the sum of discounted future excess dividends, and p^r is the sum of discounted future excess returns. We decompose the variance of excess values using $\mathrm{Var}(p - \bar{p}) = \mathrm{Var}(p^d) + \mathrm{Var}(p^r) - 2\,\mathrm{Cov}(p^r, p^d)$. The standard errors are calculated allowing for both heteroskedasticity and for the residuals to be correlated within each of the 17 years. Standard errors are in parentheses.

	Constant	r_t	$\bar{r}_t$	p_t	$\bar{p}_t$	R^2	$\frac{\mathrm{Var}(p^d)}{\mathrm{Var}(p-\bar{p})}$	$\frac{\mathrm{Var}(p^r)}{\mathrm{Var}(p-\bar{p})}$	$\frac{-2\,\mathrm{Cov}(p^r,p^d)}{\mathrm{Var}(p-\bar{p})}$
				Using q and assets weights					
r_{t+1}	0.15	0.04	−0.15	−0.11	0.01	0.03	0.54	0.18	0.28
	(0.03)	(0.03)	(0.09)	(0.02)	(0.02)		(0.09)	(0.05)	(0.09)
$\bar{r}_{t+1}$	0.16	−0.02	−0.11	−0.02	−0.02	0.03			
	(0.03)	(0.02)	(0.10)	(0.02)	(0.03)				
p_{t+1}	0.04	−0.03	−0.08	0.82	0.05	0.70			
	(0.01)	(0.03)	(0.07)	(0.02)	(0.02)				
$\bar{p}_{t+1}$	0.10	−0.01	−0.15	0.00	0.86	0.70			
	(0.02)	(0.02)	(0.09)	(0.01)	(0.03)				
				Using m and sales weights					
r_{t+1}	0.12	0.01	−0.14	−0.05	0.01	0.03	0.57	0.07	0.36
	(0.02)	(0.03)	(0.08)	(0.01)	(0.02)		(0.08)	(0.03)	(0.05)
$\bar{r}_{t+1}$	0.15	−0.02	−0.11	−0.01	−0.03	0.04			
	(0.02)	(0.01)	(0.09)	(0.01)	(0.02)				
p_{t+1}	0.01	−0.06	−0.12	0.89	0.05	0.84			
	(0.02)	(0.03)	(0.10)	(0.01)	(0.01)				
$\bar{p}_{t+1}$	0.09	−0.03	−0.19	0.03	0.85	0.80			
	(0.02)	(0.02)	(0.12)	(0.01)	(0.02)				

Table VI shows that lagged own value ratio is a strong and reliable predictor of future firm returns. The negative coefficient on lagged value ratio (−0.11 using q and −0.05 using m) reflects the value effect. Firms with high scaled prices have low subsequent returns.

Industry value ratios seem to contain little predictive information for annual firm returns, as indicated by the insignificant coefficient of 0.01 on lagged industry value ratios in the first rows of the VAR.[8] However, there is some tendency for firm value ratios to move towards industry ratios, indi-

[8] Cohen and Polk (1998) decompose book-to-market ratios into inter- and intra-industry components, and similarly find that the value effect is primarily intra-industry.

cated by the coefficient of 0.05 on lagged industry value ratios. This convergence reflects either movement of the numerator or the denominator of the firm value ratio. The numerator, market value, moves towards the industry because firm returns are slightly higher when the industry has a high value ratio (as reflected in the coefficient of 0.01). The denominator can move as well, if the firm assets (or sales) tend to move towards industry levels. Both firm value ratios and industry value ratios are strongly persistent (with coefficients of above 0.8 on own lags).

Using the coefficients of $\mathbf{A}$ defined by the regression coefficients, we calculate p^d and p^r and decompose the variance of excess values. The variance decomposition, implied by the dynamics of returns and excess values, uses the fact that $\mathrm{Var}(p - \bar{p}) = \mathrm{Var}(p^d) + \mathrm{Var}(p^r) - 2\,\mathrm{Cov}(p^r, p^d)$. Table VI shows the contribution of these three components, normalizing each component by $\mathrm{Var}(p - \bar{p})$ so that they sum to one. The scaled components of variance are

$$\frac{\mathrm{Var}(p^r)}{\mathrm{Var}(p - \bar{p})} = [\delta\Omega\delta']^{-1}\lambda\Omega\lambda' \tag{16}$$

$$\frac{\mathrm{Var}(p^d)}{\mathrm{Var}(p - \bar{p})} = [\delta\Omega\delta']^{-1}(\lambda + \delta)\Omega(\lambda + \delta)' \tag{17}$$

$$\frac{-2\,\mathrm{Cov}(p^r, p^d)}{\mathrm{Var}(p - \bar{p})} = -2[\delta\Omega\delta']^{-1}\lambda\Omega(\lambda + \delta)' \tag{18}$$

where $\lambda = (\mathbf{e}_1' - \mathbf{e}_2')(I - \rho\mathbf{A})^{-1}\mathbf{A}$, $\Omega = \mathrm{E}[\mathbf{x_t}\mathbf{x_t}'] - \mathrm{E}[\mathbf{x_t}]\mathrm{E}[\mathbf{x_t}']$, and $\delta = (\mathbf{e}_3' - \mathbf{e}_4')$ with $\mathbf{e_3}$ and $\mathbf{e_4}$ analogous to $\mathbf{e_1}$ and $\mathbf{e_2}$.

Following Campbell (1991) and Hodrick (1992), we calculate asymptotic standard errors for the variance decomposition using the delta method. We treat the VAR coefficients and the elements of Ω as parameters jointly estimated using the Generalized Methods of Moments of Hansen (1982), again allowing for heteroskedasticity and dependence within years.

The variance decomposition shows that slightly more than half (54 percent using q and 57 percent using m) of the cross-sectional variance of excess values can be explained by the differences in expected future cash flows. This fraction is significantly less than one. The remaining variation in excess values is attributable to differences in future returns and to the covariation term. The fractions of variance contributed by future returns and by covariance of returns with cash flows are each significantly different from zero. This decomposition shows the quantitative importance of predictable returns in explaining variation in excess values. If returns were totally unpredictable (so that all the coefficients in the predictive equation for returns were zero), then the procedure would mechanically attribute 100 percent of the variation to differences in future cash flows and zero to the other terms.

Another implication of the variance decomposition is that if one runs a cross-sectional regression with excess value on the left-hand side and only cash-flow related terms on the right-hand side, one should not be able to get an R-squared over 54 to 57 percent (assuming the cash flow variables used are uncorrelated with expected returns). For example, Lang and Stulz (1994) and Berger and Ofek (1995) regress excess values on size, earnings, investment, etc. They report R-squareds in the 5 to 11 percent range, so the implied upper bound is not hard to satisfy.

The covariance term (28 percent using q and 36 percent using m) is substantial. The negative correlation of p^d and p^r means that when a diversified firm has a high expected return (and thus a low excess value due to differences in returns), it also tends to have a low excess value due to differences in cash flows. Put differently, the return effect tends to magnify the cash flow effect. One could describe this covariance as consistent with "overreaction," in the sense that firms with low cash flow prospects tend to have bigger discounts than suggested by cash flows alone.

B. Regression Sample and Robustness Checks

Unlike Table II, Table VI's annual returns do not represent implementable trading strategies. First, the returns only include firms with complete returns for the entire year. Thus, both diversified and single-segment returns are subject to survivor bias. Second, the returns are from January 1 to December 31; there may be a substantial time lag between January 1 and the time a firm's data actually becomes available.

To evaluate the importance of these selection biases, we here compare the return characteristics of Tables II and VI. Calculating the annual continuously compounded total return for Table VI's sample, and comparing subsequent returns on discount firms and premium firms, produces a differential return of 3.7 percent (31 basis points per month) using q and 4.0 percent (34 basis points per month) using m. Thus the results on differential returns are similar to Panel B of Table II (which showed differentials of 29 basis points for q and 26 for m).

The homogenous VAR estimated in Table VI is obviously a gross simplification of reality. See Campbell (1991) and Hodrick (1992) for an examination of how well VARs work in the case of aggregate returns, and Vuolteenaho (1999) for an examination of cross-sectional VARs similar to ours. Given the traditional emphasis on medians in the diversification literature, one might also worry that outliers heavily influence our results. One way of assessing the ability of our simple model to represent reality is to see whether it can match important characteristics of the data. Using the results from the first two regressions in Table VI, we form annual forecasts for differential returns for discount firms and for premium firms. We find that the forecasts closely match the realized differential returns: The forecast is 4.0 percent using q and 3.8 percent using m (similar to the actual differentials of 3.7 and 4.0 percent, previously mentioned).

We now report further robustness checks on Table VI. First, Fama–MacBeth (1973) estimation produces regression results similar to the ones in Table VI, with the fraction of variance attributable to $\mathrm{Var}(p^d)$ rising to

0.64 for q and 0.60 for m.[9] Second, we tried dropping extreme values, defined as any observation in which any one of the eight current or lagged variables was in the top or bottom five percent of its distribution. Using this sample (about half as large as the baseline sample) produces similar results, with the fraction of variance attributable to $\mathrm{Var}(p^d)$ of 0.58 for q and 0.67 for m. Third, using excess returns (by subtracting continuously compounded T-bill returns from total returns) instead of total returns also produces similar results, with the fraction of variance staying at 0.54 for q and 0.57 for m.

The results from Table II show that premium firms have negative excess returns but discount firms have excess returns of about zero. Based on these results, one might wonder if cash flows are more important for discount firms than for premium firms. We therefore split the sample into premium firms and discount firms (as of year $t - 1$) and calculated the variance decomposition for these two groups, with separate VARs. For q, the fraction of variance attributable to $\mathrm{Var}(p^d)$ was 0.41 for premium firms and 0.82 for discount firms; for m, the fraction was 0.71 for premium firms and 0.60 for discount firms. Thus there is no consistent pattern in the relative importance of cash flows for premium and discount firms.

The homogenous model of Table VI forces all firms to have the same coefficients. This constraint implies that all firms have the same long-run value ratio, for example, and does not allow different firms to have permanently different expected returns or different expected cash flows. An alternative estimation strategy is to allow firm-specific fixed effects by differencing the variables in Table VI. A previous version of this paper, Lamont and Polk (1999), estimates a first differenced version of Table VI and finds similar results, with an attribution to $\mathrm{Var}(p^d)$ of 0.60 for q and 0.42 for m.

In summary, we have no reason to believe that either selection biases or outliers are quantitatively important for our variance decomposition. While different methods produce somewhat different estimates, all estimates of the fraction of cross-sectional variance attributable only to cash flows are less than one. The estimates range from 0.42 to 0.75.

C. Is Anything Special About Diversified Firm Value Ratios?

As discussed previously, one explanation of our results is simply that the value effect is present in diversified firm stock returns. This explanation leads naturally to the question of whether the cross-section of diversified firm value ratios is in any way different from the cross section of single-segment firm value ratios. Single-segment firms also have value ratios that (for individual firms) are not always identical to the value ratios of their

[9] Stambaugh (1999) discusses a small sample bias in time-series predictive regressions of returns on lagged scaled values. Since our regression has a time-series dimension as well as a cross-sectional dimension, it is subject to this bias. Since the pooled OLS results are so similar to the Fama-MacBeth (1973) results (which are based on purely cross-sectional regressions with no time-series dimension), the bias is unimportant in our sample.

matching portfolio. These differences again must be due to either differences in future returns or in future cash flows. Do the sources of variation in industry-adjusted value ratios look the same for single-segment firms?

To answer this question, we formed excess value ratios and excess returns for single-segment firms. For each single-segment firm, we form a benchmark portfolio of other single-segment firms in the same industry, using the same matching and weighting algorithm as before. When forming the industry benchmark, we exclude the target firm from the set of possible matching firms. The resulting sample of firms is larger than the sample of diversified firms, with about 22,000 observations on firms that meet the VAR's data requirements.

Table VII shows results from a vector autoregression using single-segment firms instead of diversified firms. Looking first at the regression coefficients, the results are quite similar to Table VI. Like diversified firms, single-segment firm returns are negatively related to their value ratio. Like diversified firms, single-segment firms have excess value ratios that are highly persistent and that have a slight tendency to converge towards industry benchmark levels.

Looking at the variance decomposition, the comparison is slightly more ambiguous. For m, the variance decomposition for single-segment firms looks quite similar to the variance decomposition for diversified firms, with the fraction of variance attributable to $\mathrm{Var}(p^d)$ at 67 percent. For q, the results are somewhat different. The fraction of variance attributable to $\mathrm{Var}(p^d)$ is 89 percent, and one cannot reject the hypothesis that the fraction is 100 percent.

On the other hand, there certainly is some predictability of returns and that predictability creates sizable and statistically significant variance of p^r that happens to be offset by the covariance term. And the confidence interval for variance attributable to $\mathrm{Var}(p^d)$ for q goes down to 0.49, so the q results in Table VII do not present strong evidence that single-segment firms are different from diversified firms. The standard errors for the q results in Table VII are strikingly large, particularly for the covariance term, suggesting that the covariance estimate is unreliable.

These findings suggest there is nothing special about the cross section of diversified firms' returns and value ratios. The effect we find, that excess values are negatively correlated with subsequent returns, simply reflects the well-known value effect in stocks. Stocks with high scaled prices have low subsequent returns, and this holds true for both single-segment and multisegment firms.

VI. Conclusions

A. Summary

We show that firms with diversification discounts have high subsequent returns and firms with premia have low subsequent returns. This pattern in diversified firm returns is a manifestation of the familiar value effect,

Table VII

Dynamic Behavior of Annual Returns and Values Ratios for Single-segment Firms and Matching Portfolios, 1981 to 1997

Vector autoregression results using pooled OLS estimation. The regression is $\mathbf{x_{t+1}} = \mathbf{c} + \mathbf{A}\mathbf{x_t} + \mathbf{e_{t+1}}$ where $\mathbf{x_t'} = [r_t, \bar{r}_t, p_t, \bar{p}_t]$ is the vector of continuously compounded returns and log value ratios. The sample is 22,015 single-segment firm-years that have excess values and excess returns for two consecutive years. Using the estimated coefficients, we forecast future returns for each firm and calculate p^d and p^r. p^d is a function of the sum of discounted future excess dividends, and p^r is the sum of discounted future excess returns. We decompose the variance of excess values using $\text{Var}(p - \bar{p}) = \text{Var}(p^d) + \text{Var}(p^r) - 2\,\text{Cov}(p^r, p^d)$. The standard errors are calculated allowing for both heteroskedasticity and for the residuals to be correlated within each of the 17 years. Standard errors are in parentheses.

	Constant	r_t	$\bar{r}_t$	p_t	$\bar{p}_t$	R^2	$\frac{\text{Var}(p^d)}{\text{Var}(p-\bar{p})}$	$\frac{\text{Var}(p^r)}{\text{Var}(p-\bar{p})}$	$\frac{-2\,\text{Cov}(p^r,p^d)}{\text{Var}(p-\bar{p})}$
				Using q and assets weights					
r_{t+1}	0.13	0.05	−0.01	−0.08	−0.04	0.02	0.89	0.29	−0.18
	(0.03)	(0.02)	(0.07)	(0.02)	(0.02)		(0.20)	(0.13)	(0.31)
$\bar{r}_{t+1}$	0.14	−0.01	−0.01	0.00	−0.01	0.00			
	(0.03)	(0.01)	(0.08)	(0.01)	(0.03)				
p_{t+1}	0.05	−0.05	0.02	0.82	0.02	0.69			
	(0.02)	(0.02)	(0.05)	(0.02)	(0.02)				
$\bar{p}_{t+1}$	0.09	−0.03	−0.06	0.02	0.88	0.74			
	(0.02)	(0.01)	(0.08)	(0.01)	(0.03)				
				Using m and sales weights					
r_{t+1}	0.08	0.04	−0.05	−0.04	0.00	0.01	0.67	0.12	0.22
	(0.03)	(0.02)	(0.07)	(0.01)	(0.01)		(0.07)	(0.04)	(0.07)
$\bar{r}_{t+1}$	0.14	−0.01	0.00	0.00	−0.02	0.01			
	(0.03)	(0.01)	(0.07)	(0.00)	(0.02)				
p_{t+1}	−0.03	−0.04	−0.02	0.88	0.05	0.83			
	(0.03)	(0.02)	(0.07)	(0.01)	(0.01)				
$\bar{p}_{t+1}$	0.07	−0.04	−0.09	0.03	0.90	0.85			
	(0.03)	(0.01)	(0.08)	(0.00)	(0.02)				

previously documented in the cross section of average equity returns for all firms. Current asset valuations are negatively related to future returns.

Since actual returns are consistently higher for discount firms than premium firms, we argue that expected returns are also higher for discount firms than premium firms. The pattern of returns does not appear to only reflect surprises or news that happened to occur in the sample period. Thus the diversification puzzle is both an expected return phenomenon and an expected cash flow phenomenon.

Using simple present value relations and a first-order vector autoregression, we estimate the fraction of the cross-sectional variance of excess values that can be attributed to differences in future cash flows. We find that slightly more than half of the variance is due to future cash flow differences between diversified firms and single-segment firms, with the remaining half due to differences in future returns and the covariance between returns and cash flows.

B. Interpretation

Our analysis has limitations. Our variance decomposition is only as accurate as our statistical model of returns, which is undoubtedly incomplete since it forecasts returns with a small set of variables. Although we show time-series evidence suggesting that an unexpected surge in corporate control activity is not driving our results, these unexpected events do affect our estimates. Thus our point estimate, that pure cash flow effects can account for 54 percent of the cross-sectional variance, may understate or overstate the true fraction.

We report good news for those who believe that diversification destroys value. Our estimates leave plenty of room for wasteful spending or cross-subsidization to reduce cash flow, since we find a large role for variation in future cash flow. And, of course, returns are only one part of the story for diversified firms, and equally valuable is a large body of existing evidence on earnings, capital expenditures, and productivity.

The main message of this paper is simply that the value effect exists among diversified firms, so that one cannot treat value ratios as measuring future cash flows only. How one interprets the results depends on one's view of the value effect. For believers in rational risk-based asset pricing, the interpretation is that premium firms have low discount rates, and discount firms have high discount rates. Under this view, diversification does not destroy value (unless it somehow increases risk).

For believers in behavioral finance, the value effect reflects mispricing by irrational investors. This belief suggests that the diversification discount reflects (at least partially) irrationally low prices. A specific version of this story is that diversified firms are complex, and so investors and analysts have difficulty valuing diversified firms (see the discussion in Hadlock, Ryngaert, and Thomas (1999) for example). This difficulty makes diversified stocks "neglected," misunderstood, and undervalued. According to this view, diversification does indeed destroy value, in the sense that breaking up the firm into its components would raise the combined valuation.

The mispricing view also can explain the finding of Berger and Ofek (1996) that discount firms tend to be taken over. The traditional interpretation of these takeovers (based on cash flows) is that discount firms are poorly managed and can be improved by better management or better incentives. The alternative interpretation is that the discount firms are undervalued, and therefore represent a window of opportunity for smart investors to take over the firm at a bargain price. This alternative view does not imply that diversification necessarily destroys value, since Berger and Ofek (1996) find that focused firms with a discount are also more likely to be taken over.

Since our results are based on the cross-sectional variation in excess values, they say nothing about why the average diversified firm is worth less than the sum of its parts. Nevertheless, one can speculate that the same effect that explains part of the deviation from average might also explain part of the average.

Appendix

A. Data Sources and Definitions

Our data on segments comes from several Current and Research segment files obtained from Wharton Research Data Services in April 1999. Our firm-level COMPUSTAT and CRSP data comes from the University of Chicago's CRSP, in August 1999. Total returns on the Lehman Brothers Corporate Bond Index are provided by Ibbotson Associates. In our calculation of market value, we use CRSP market equity.

We define Q as {market capitalization (from CRSP) + book assets (data item 6) − book equity (data item 60) − deferred taxes (data item 74)}/book assets (data item 6). We define leverage as total debt/(total debt + market capitalization) where total debt is defined as long-term debt (data item 9) + debt in current liabilities (data item 34) + redemption value of preferred stock (data item 56). We define M as (total debt + market capitalization)/net sales (data item 12).

In some cases, CRSP recorded delisting prices several months after the security ceased trading and thus after a period of missing returns. In these cases, we calculated the total return from the last available price to the delisting price, and prorated this return over the intervening months.

For firms with multiple classes of stock, in calculating market equity and stock returns, we aggregate all separate classes of stock together into one value-weighted portfolio.

B. Screening

We drop firm-years if any of the following conditions hold: it has missing or nonpositive firm sales or firm assets; has missing or nonpositive (for any segment) segment sales or segment assets; has any segment that COMPUSTAT assigns a 1-digit SIC code of 0, 6, or 9; the sum of segment sales is not within one percent of the total sales of the firm; the firm changes the month of its fiscal year-end such that in December of year $t-1$ our latest information is from year $t-2$. We also drop firms (such as GM) who report multiple firm totals for the same year (that is, firms which report different COMPUSTAT total sales for the same CRSP permanent company identifier number).

When calculating monthly returns, we also impose a constraint to deal with COMPUSTAT backfilling (a practice which may induce survivor bias). We require that firms have at least two years of COMPUSTAT data prior to year t.

REFERENCES

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid ask spread, *Journal of Financial Economics* 17, 233–249.

Ball, Ray, 1978, Anomalies in relationships between securities' yields and yield-surrogates, *Journal of Financial Economics* 6, 103–126.

Berger, Philip G., and Eli Ofek, 1995, Diversification's effect on firm value, *Journal of Financial Economics* 37, 39–66.

Berger, Philip G., and Eli Ofek, 1996, Bustup takeovers of value-destroying diversified firms, *Journal of Finance* 51, 1175–1200.

Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.

Campbell, John Y., and Robert J. Shiller, 1988, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661–676.

Capozza, Denis R., and Paul J. Seguin, 1999, Focus, transparency and value: The REIT evidence, *Real Estate Economics* 27, 587–619.

Chevalier, Judith, 1999, Why do firms undertake diversifying mergers? An examination of the investment policies of merging firms, Working paper, http://gsbwww.uchicago.edu/fac/judith.chevalier/research/

Cohen, Randolph B. and Christopher Polk, 1998, The impact of industry factors in stock returns, Working paper, http://www.kellogg.nwu.edu/faculty/polk/research/

Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho, 2000, The value spread, Working paper, http://www.kellogg.nwu.edu/faculty/polk/research/

Comment, Robert, and Gregg A. Jarrell, 1995, Corporate focus and stock returns, *Journal of Financial Economics* 37, 67–87.

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389–406.

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

Elton, Edwin J., 1999, Expected return, realized return, and asset pricing tests, *Journal of Finance* 54, 1199–1220.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Gatev, Evan G., William N. Goetzmann, and K. Geert Rouwenhorst, 1999, Pairs trading: performance of a relative value arbitrage rule, Working paper, Yale School of Management.

Hadlock, Charles, Michael Ryngaert, and Shawn Thomas, 1999, Corporate structure and equity offerings: Are there benefits to diversification? Working paper, Michigan State University.

Hansen, Lars Peter, 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.

Hecht, Peter, 1999, The cross section of expected firm, not equity, returns, Working paper, Harvard University.

Hodrick, Robert J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357–386.

Lamont, Owen, 1997, Cash flow and investment: Evidence from internal capital markets, *Journal of Finance* 52, 83–109.

Lamont, Owen, and Christopher Polk, 1999, The diversification discount: Cash flows vs. returns, NBER Working Paper 7396.

Lang, Larry, and René Stulz, 1994, Tobin's Q, corporate diversification, and firm performance, *Journal of Political Economy* 102, 1248–1280.

Maksimovic, Vojislav and Gordon Phillips, 1999, Do conglomerate firms allocate resources inefficiently? Working paper, http://www.rhsmith.umd.edu/Finance/gphillips/

Morck, R., A. Shleifer, and R. Vishny, 1990, Do managerial objectives drive bad acquisitions? *Journal of Finance* 45, 31–48.

Pontiff, Jeffrey, 1996, Costly arbitrage: Evidence from closed-end funds, *Quarterly Journal of Economics* CXI, 1135–1151.

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

Rogers, William, 1993, Regression standard errors in clustered samples, *Stata Technical Bulletin* 13, 19–23.

Scharfstein, David S., 1998, The dark side of internal capital markets II, NBER Working Paper No. 6352.

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

Servaes, Henri, 1996, The value of diversification during the conglomerate merger wave, *Journal of Finance* 51, 1201–1225.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35–55.

Shumway, Tyler, 1997, The delisting bias in CRSP data, *Journal of Finance* 52, 327–340.

Stambaugh, Robert F., 1999, Predictive regressions, NBER Technical Working Paper 240.

Vuolteenaho, Tuomo, 1999, What drives firm level stock returns? Working paper, Harvard University.