PORTFOLIO
MANAGEMENT
RESEARCH
By With Intelligence

the journal of PORTFOLIO management

volume 46
number 1

NOVEMBER 2019

jpm.pm-research.com

Why Do Enterprise Multiples Predict Expected Stock Returns?

Steven S. Crawford, Wesley R. Gray, and Jack Vogel

alpha architect

Founded in 2010, the Alpha Architect mission is to empower investors through education. We are a research-intensive asset management firm with a focus on high-conviction value and momentum factor exposures.

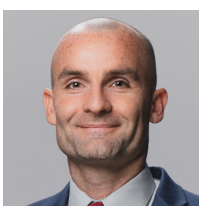
We seek to deliver "Affordable Alpha," which means highly differentiated investment strategies at lower costs, thereby giving sophisticated investors a higher chance of winning, net of fees and taxes. We currently offer Exchange-Traded Funds (ETFs), Separately Managed Accounts (SMAs), and ETF White-label services.

Our core beliefs are as follows: 1) Transparency, 2) Evidence-Based Investing, 3) Systematic Decision Making, and 4) Win-Win client relationships.

The firm is based in the suburbs of Philadelphia, PA (Havertown, PA). We are a service-disabled and minority-owned firm.

Steven S. Crawford

Steve Crawford is an associate professor of Accounting at the C. T. Bauer College of Business. He teaches Financial Statement Analysis in the MSACCY and MBA programs and has won numerous teaching awards. Steve has research interests in financial analysts, financial markets, disclosure, and compensation. His research papers have been published in leading accounting and finance journals. Steve earned his MBA and a Ph.D. from the University of Chicago's Booth School of Business in 2007. He received his BA and M.Acc. from Brigham Young University in 2002.



Wesley R. Gray

After serving as a Captain in the United States Marine Corps, Wes earned an MBA and a PhD in finance from the University of Chicago where he studied under Nobel Prize Winner Eugene Fama. Next, Wes took an academic job in his wife's hometown of Philadelphia and worked as a finance professor at Drexel University. Wes's interest in bridging the research gap between academia and industry led him to found Alpha Architect, an asset management firm dedicated to an impact mission of empowering investors through education. He is the CEO and co-CIO of Alpha Architect. He is a contributor to multiple industry publications and regularly speaks to professional investor groups across the country. Wes has published multiple academic papers and four books, including Embedded (Naval Institute Press, 2009), Quantitative Value (Wiley, 2012), DIY Financial Advisor (Wiley, 2015), and Quantitative Momentum (Wiley, 2016). Wes currently resides in Palmas Del Mar Puerto Rico, with his wife and three children.

Jack Vogel

Jack Vogel is the co-CIO and CFO of Alpha Architect. Jack conducts research in empirical asset pricing and behavioral finance. Jack is a co-author of DIY FINANCIAL ADVISOR: A Simple Solution to Build and Protect Your Wealth and QUANTITATIVE MOMENTUM: A Practitioner's Guide to Building a Momentum-Based Stock Selection System. His academic background includes experience as an instructor and research assistant at Drexel University in both the Finance and Mathematics departments and a Finance instructor at Villanova University. He has a Ph.D. in Finance and an MS in Mathematics from Drexel University. Jack graduated summa cum laude with a BS in Mathematics and Education from The University of Scranton.

Why Do Enterprise Multiples Predict Expected Stock Returns?

Steven S. Crawford, Wesley R. Gray, and Jack Vogel

Steven S. Crawford
is an assistant professor of accounting at the University of Houston in Houston, TX.
scrawford@bauer.uh.edu

Wesley R. Gray
is the chief executive officer of Alpha Architect in Broomall, PA.
wes@alphaarchitect.com

Jack Vogel
is the chief investment officer/chief financial officer of Alpha Architect in Broomall, PA.
jack@alphaarchitect.com

KEY FINDINGS

- We revisit the enterprise multiple (EM) effect and document that the EM effect is primarily attributable to mispricing and cannot be explained by higher systematic risk.
- We document that the EM effect is stronger during times of strong market sentiment, which is further evidence that the effect is driven by mispricing.
- Over 80% of the alpha associated with the best EM portfolio is generated by the short leg. If managing short positions is costly, these results suggest that the mispricing associated with the high-mispricing EM portfolio is difficult to exploit profitably.

ABSTRACT: *The enterprise multiple (EM) effect has been documented across global stock markets. EM is a robust predictor of expected average returns and generates a stronger value effect than traditional value metrics. We find evidence that the EM effect is primarily attributable to mispricing and cannot be explained by higher systematic risk. We document that earnings announcement returns, forecast errors, and forecast revisions all support the notion that the EM effect is driven by mispricing associated with predictable investor expectation errors. Finally, we show that the EM effect is stronger during times of strong market sentiment, which also supports the mispricing-based hypothesis.*

TOPICS: ***Factor-based models, equity portfolio management, portfolio construction, portfolio theory****

*All articles are now categorized by topics and subtopics. **View at PM-Research.com.**

Loughran and Wellman (2011) documented the enterprise multiple (EM) effect, where the EM is calculated as enterprise value (EV = Equity + Debt + Preferred stock—Cash) divided by earnings before interest, taxes, depreciation, and amortization (EBITDA; operating income before depreciation and amortization[1]). The EM effect can be described as the ability of EM to predict the cross section of expected returns better than alternative valuation metrics (e.g., book-to-market [B/M]) traditionally used in the literature. Walkshäusl and Lobe (2015) extended the analysis of Loughran and Wellman (2011) to international markets and found that the EM effect is even stronger outside of the United States.

We contribute to the literature by investigating *why* EM is such a powerful predictor of expected average returns relative to traditional value measures such as B/M.

[1] We conduct all the analysis with operating income after depreciation and amortization (or EBIT) and find quantitatively similar results. Gray and Carlisle (2012) and Gray and Vogel (2012) also documented an EM effect.

A key question in asset pricing is whether high average expected returns associated with value stocks and low average expected returns earned by glamour stocks are compensation for risk or a result of systematic mispricing. The debate is mixed over the source of value stock returns, with some researchers suggesting that the excess returns are due to risk (e.g., the risk-based hypothesis presented by Fama and French 1993) and other authors claiming that the average returns are due to mispricing (e.g., the mispricing-based hypothesis presented by Lakonishok, Shleifer, and Vishny 1994). The more powerful EM value effect serves only to reinvigorate the debate and provides a new opportunity to better understand the value anomaly.

To better understand the EM effect, we use the empirical framework presented by Piotroski and So (2012) to differentiate between the risk-based and the mispricing-based hypotheses. The authors create two test portfolios: (1) long value firms and short glamour firms with high investor expectation errors and (2) long value firms and short glamour firms with low investor expectation errors. The core empirical tests examine the spread between the two portfolios. The risk-based hypothesis predicts no difference between the returns of these portfolios, and the mispricing-based hypothesis predicts a positive spread in returns.

In our analysis, we create test portfolios similar to those used by Piotroski and So (2012) by sorting stocks on EM and 11 proxies for the fundamental value of the stock. We find strong evidence in favor of the mispricing-based hypothesis and weak evidence that the EM effect is a proxy for higher discount rates, as alluded to by Loughran and Wellman (2011). Specifically, portfolios with high investor expectation errors earn higher returns than portfolios with low investor expectation errors. We also examine earnings announcement returns, forecast errors, and forecast revisions for our test portfolios. The evidence from this analysis supports the notion that the EM effect is driven, at least in part, by mispricing associated with predictable investor expectation errors. Finally, we show that the EM effect is larger during times of strong market sentiment, which also supports the mispricing-based hypothesis.

We perform a battery of robustness tests on our core results. We break our sample into several time periods and find little difference across samples. We test for calendar effects by eliminating all January months and find little change to our results. Finally, we explore a variety of asset pricing models and find no evidence to suggest that our core results are driven by a particular test we use in our analysis.

If, however, the EM effect can be attributed to a mispricing phenomenon (at least in part), why have market participants not eliminated the opportunity? To address this question we examine the limits to arbitrage associated with exploiting the best EM portfolio strategy. We examine the long and short legs of the strategy and show that over 80% of the alpha associated with the best EM portfolio is generated by the short leg. To the extent that managing short positions is costly, these results suggest that the mispricing associated with the high-mispricing EM portfolio is difficult to profitably exploit. In addition, if costly market frictions continue to exist and investor expectation errors persist, we can expect that the EM effect may continue in the future.

HYPOTHESIS DEVELOPMENT

A key debate in asset pricing is whether the higher average returns associated with value stocks (e.g., high B/M stocks) and the lower average returns earned by glamour stocks (e.g., low B/M) are compensation for risk or due to systematic mispricing. Fama and French (1993) argued that B/M is a proxy for unobserved risk factors. A related literature offers evidence to support this conjecture by showing that value (high B/M) and glamour (low B/M) firms covary differently with macroeconomic risks in the economy (e.g., Campbell, Polk, and Vuolteenaho 2010; Santos and Veronesi 2010). However, Lakonishok, Shleifer, and Vishny (1994) set the stage for an alternative to the risk-based argument by presenting evidence that value stocks earn higher returns relative to glamour stocks because investors make systematic errors in their expectations about the future profitability of extreme B/M firms. In other words, B/M (or similar price-based ratios) identifies mispricing, not risk.

The debate over whether the value/glamour anomaly is due to systematic mispricing or compensation for risk continues in the literature. In an attempt to shed light on the issue, Piotroski and So (2012) tested the mispricing hypothesis versus the risk-based theory with the Piotroski (2000) F-score. The F-score consists of nine accounting signals related to a firm's fundamentals, which Piotroski and So (2012) used to proxy for a firm's fundamental value.[2] Specifically, the F-score is used to separate firms in glamour and value portfolios into firms with low and high fundamental value. The combination of the value/glamour sorts and fundamental value sorts allows them to proxy for expectation errors. For example, value firms with high F-scores are assumed to have high investor expectation errors: Value firms are expected to perform poorly, but firms with high F-scores have been shown empirically to have high future profitability and earnings (high fundamental value). Similarly, glamour firms with low F-scores have high expectation errors because glamour firms are expected to perform well, but they have lower future profitability and earnings (low fundamental value). Alternatively, value (glamour) firms with low (high) F-scores are assumed to have low expectation errors because they are expected to perform poorly (well) and they have weak (strong) expected fundamentals.

The authors create a high-mispricing portfolio that goes long value firms with high expectation errors (high B/M, high F-score) and short glamour firms with high expectation errors (low B/M, low F-score). Their low-mispricing strategy goes long value firms with low expectation errors (high B/M, low F-score) and short glamour firms with low expectation errors (low B/M, high F-score). Piotroski and So (2012) considered the high-mispricing strategy to be a portfolio that captures the highest degree of mispricing and the low-mispricing strategy to be a portfolio that captures the least amount of mispricing. The mispricing hypothesis suggests that the high-mispricing portfolio will produce positive abnormal returns, and the low-mispricing portfolio will not.

We leverage the research design of Piostroski and So (2012) to ascertain whether mispricing or risk drives the EM effect. We create high-mispricing and low-mispricing long–short EM-sorted portfolios and conduct performance analysis on the two strategies. Strong performance for the high-mispricing portfolio and weak performance for the low-mispricing portfolio suggests systematic mispricing that is unexplained by our traditional notions of risk. Because the F-score is described by Piotroski (2000) as somewhat "ad-hoc," we use 11 more conventionally accepted variables identified by Stambaugh, Yu, and Yuan (2012), in conjunction with the EM sorts, to measure ex ante fundamental value. These variables predict differences in the cross section of average expected returns that exist after accounting for risk adjustment models such as the three-factor model described by Fama and French (1993). We use these 11 measures (and a combo measure) to proxy for the fundamental value of firms.[3]

DATA

To execute our empirical analyses, we use a large sample of firms traded on the major stock exchanges (New York Stock Exchange [NYSE], American Stock Exchange [AMEX], and NASDAQ). To test whether the EM effect is due to mispricing or risk, we then obtain necessary accounting information from Compustat and market data (returns, market value of equity) from CRSP. Consistent with Loughran and Wellman (2011), we limit our sample to firms with ordinary common equity on CRSP and eliminate all real estate investment trusts, American depositary receipts, closed-end funds, and financial firms. We also exclude firms with negative book values, firms with negative EV, and firms with negative EBITDA values. We incorporate CRSP delisting return data using the technique developed by Beaver, McNichols, and Price (2007). To be included in the sample, all firms must have a nonzero market value of equity as of June 30 of year *t*.

Our main tests focus on examining returns in portfolios double sorted on the EM and several measures of fundamental value used by Stambaugh, Yu, and Yuan (2012). We measure the EM as EV divided by EBITDA. We follow Loughran and Wellman (2011) and calculate EV as the market value of equity calculated from CRSP plus total debt (Compustat data items DLC and DLTT [short- and long-term debt]) plus preferred stock value (item PSTKRV) minus cash and short-term investments (item CHE). We define EBITDA as operating income before depreciation (Compustat item OIBDP).

[2] The F-score is built with nine 0/1 indicators that are summed to give each firm a score between 0 and 9. An F-score of 0 is the worst, whereas an F-score of 9 is the best.

[3] Stambaugh, Yu, and Yuan (2012) referred to their 12 variables as anomalies. In this article, we use these anomalies to proxy for the fundamental value of the firm given that these proxies predict future returns after accounting for traditional risk metrics. The variables overlap with many of the variables in the F-score. We also examine the F-score and find similar results.

Stambaugh, Yu, and Yuan (2012) identified 11 well-documented anomalies that serve as proxies for a firm's fundamental value in our analysis and which we use to determine whether a specific EM portfolio has high or low expectation errors. We now briefly describe the variables from their paper.

Financial distress (DISTRESS) is computed using the methodology of Campbell, Hilscher, and Szilagyi (2008), who found that firms with high failure probability have lower subsequent returns. Their methodology involves estimating a dynamic logit model with both accounting and equity market variables as explanatory variables. The variables included in their model are net income divided by the market value of assets, total liabilities divided by the market value of assets, excess stock return, the volatility of a firm's stock returns, firm size measured relative to the market, share price, the market-to-book ratio, and cash and short-term assets divided by the market value of assets.

The O-Score (OSCORE) is a conditional logit model developed by Ohlson (1980) to calculate the probability of bankruptcy. The variables included in the model are the log of total assets, total liabilities divided by total assets, working capital divided by total assets, current liabilities divided by current assets, net income divided by total assets, funds provided by operations divided by total liabilities, and the change in net income. The model also included two indicator variables: (1) a variable set to one if total liabilities exceed total assets and zero otherwise; (2) a variable set to one if net income was negative for the last two years and zero otherwise.

We measure net stock issuance (NETISS) as the growth rate of the split-adjusted shares outstanding in the previous fiscal year. Ritter (1991) and Loughran and Ritter (1995) showed that, in post-issue years, equity issuers underperform matching nonissuers with similar characteristics. The evidence suggests that investors are unable to recognize that firms prefer to raise capital by issuing stock when equity prices are overvalued.

We follow Daniel and Titman (2006) and measure composite equity issuance (COMPISS) as the amount of equity a firm issues (or retires) in exchange for cash or services. In other words, this measure captures the change in a firm's market value that is not attributable to stock returns. Daniel and Titman (2006) found that issuers underperform nonissuers because investors overlook the signals from repurchases and issuance.

We measure total accruals (ACCRUALS) following Sloan (1996). Specifically, accruals are calculated as the change in current assets (less the change in cash) minus the change in current liabilities (less the change in short-term debt and the change in income taxes payable) minus depreciation and amortization expense. Sloan (1996) found that firms with high accruals earn abnormal lower returns on average than firms with low accruals and reasoned that investors overestimate the persistence of the accrual component of earnings vis-à-vis the cash component of earnings.

Net operating assets (NOA) is computed following the methodology of Hirshleifer et al. (2004) as operating assets minus operating liabilities scaled by total assets. Hirshleifer et al. (2004) found that NOA is a strong negative predictor of long-run stock returns. The reasoning behind the anomaly is similar to the accrual anomaly: Investors focus on accounting profitability while neglecting information about cash profitability.

The momentum effect was first discovered by Jegadeesh and Titman (1993). We calculate the momentum monthly return (MOM) by looking at the cumulative returns from month -12 to month -2, similar to Fama and French (2008).

The gross profitability premium (GP) was first documented by Novy-Marx (2013), who showed that sorting on gross profit-to-assets (sales [SALE] minus cost of goods sold [COGS], scaled by total assets [AT]) creates abnormal benchmark-adjusted returns, with more profitable firms having higher returns than less profitable ones. Novy-Marx (2013) argued that gross profits divided by AT is the cleanest accounting measure of true economic profitability and that investors overlook the investment value of the profitability of the firm.

We measure asset growth (AG) as the growth rate of the AT (item AT) in the previous fiscal year. Cooper, Gulen, and Schill (2008) found that firms with high AG earn lower subsequent returns relative to firms with low AG. The authors argued that this return pattern is driven by investors overestimating future growth and business prospects based on observing a firm's AG.

We calculate return on assets (ROA) as income before extraordinary items (IB) divided by AT. In two related papers, Fama and French (2006) and Chen, Novy-Marx, and Zhang (2010) documented that investors appear to underestimate the importance of ROA in explaining future returns. Namely, firms with high (low) prior ROA earn abnormally high (low) returns.

We measure investment-to-assets (INV) as the annual change in gross property, plant, and equipment (PPEGT) plus the annual change in inventories (INVT) scaled by the lagged book value of assets (AT). Titman, Wei, and Xie (2004) and Xing (2008) found that higher past investment predicts abnormally lower future returns. The authors posited that this anomaly stems from investors' inability to identify manager empire-building behavior via investment patterns.

We create a combination metric (COMBO) by calculating ranks for each of the 11 variables identified by Stambaugh, Yu, and Yuan (2012) and then averaging those ranks across all the variables.[4] Exhibit 1 gives the summary statistics for each of the 11 fundamental value proxies described. These values are winsorized at the 1% and 99% level each year to eliminate outliers. Panel B (low EM value firms) and Panel C (high EM glamour firms) highlight that the fundamental value proxies associated with glamour firms have a higher standard deviation than value firms, with the one exception being DISTRESS. The value and glamour firm means and medians are similar for DISTRESS, OSCORE, and ACCRUAL, whereas the other eight measures differ across value and glamour firms.

PERFORMANCE ANALYSIS

High-Mispricing and Low-Mispricing Portfolio Performance

For our main empirical analysis, we conduct a double sort of our sample firms into value and glamour quintiles based on the firms' EMs and then sequentially sort these EM portfolios into quintiles based on each of the 12 fundamental value proxies (i.e., the 11 individual proxies and the combination proxy). Value (low EM) and glamour (high EM) firms are identified using the 20th and 80th NYSE EM cutoffs. Firms with high and low fundamental values are identified using the 20th and 80th percentile cutoffs of each of the 11 fundamental value proxies. We isolate a high-mispricing portfolio, which goes long value firms (low EM) with high expectation errors (cheap with high fundamental value) and goes short glamour firms (high EM) with high expectation errors (expensive with low fundamental value). We also create a low-mispricing portfolio, which goes long value firms (low EM) with low expectation errors (cheap with low fundamental value) and short glamour firms (high EM) with low expectation errors (expensive with high fundamental value). The sample uses information available in June of year t to forecast the returns from July of t to June of year $t + 1$. The exception is the momentum variable, which is measured each month to maintain continuity with prior research. We then calculate monthly value-weighted portfolio returns[5] and analyze the return series using the calendar-time portfolio regression approach because of the statistical problems inherent in long-run buy-and-hold abnormal returns, as discussed by Mitchell and Stafford (2000) and Fama (1998).

The *capital asset pricing model (CAPM) alpha* is a risk-adjusted return equal to the intercept from a time-series regression of the long–short portfolios on the excess return on the value-weight market index. The *three-factor alpha* is a risk-adjusted return equal to the intercept from a time-series regression of the long–short portfolios on the excess return on the value-weight market index, the return on the size (SMB) factor, and the return on the value (HML) factor (see Fama and French 1993). The *four-factor alpha* is a risk-adjusted return equal to the intercept from a time-series regression of long–short portfolios on the excess return on the value-weight market index, the return on SMB, the return on HML, and the return on a prior-year return momentum (MOM) factor.

Exhibit 2 presents the estimated alphas from our calendar-time portfolio regressions. The estimates in Exhibit 2 represent the mean monthly abnormal return over the calendar-time horizon for the low-mispricing (Panel A) and high-mispricing portfolios (Panel B) using the different proxies for a firm's fundamental value. In the low-mispricing portfolios, the reported four-factor alpha estimates are not statistically different from zero. In contrast, the four-factor alpha estimates for the high-mispricing portfolios are positive and significant at the 5% level in every instance. Furthermore, we report a paired t-test for differences in the average returns of the low-mispricing and high-mispricing portfolios for all 12 fundamental value proxies

[4] To ensure that all variables are aligned (higher = better, lower = worse), we invert some measures.

[5] Findings using equal-weighted results are similar, albeit more extreme.

Exhibit 1
Summary Statistics for Fundamental Value Proxies

	DISTRESS	OSCORE	NETISS	COMPISS	ACCRUAL	NOA	MOM	GP	AG	ROA	INV
Panel A: All Firms											
N	128,813	126,075	129,507	127,992	125,926	129,261	127,992	129,507	129,506	129,506	127,457
Mean	0.105	0.045	0.040	0.013	–0.021	0.736	0.170	0.412	0.197	0.064	0.108
Median	0.001	0.016	0.004	0.000	–0.030	0.720	0.080	0.370	0.093	0.055	0.066
Standard Deviation	0.264	0.081	0.146	0.112	0.091	0.317	0.571	0.248	0.439	0.085	0.191
25th Percentile	0.000	0.005	0.000	–0.032	–0.069	0.577	–0.172	0.227	0.009	0.022	0.015
75th Percentile	0.016	0.045	0.025	0.018	0.017	0.852	0.375	0.542	0.225	0.100	0.149
Panel B: Value Firms											
N	30,745	30,745	30,745	30,745	30,745	30,745	30,745	30,745	30,745	30,745	30,745
Mean	0.144	0.034	0.013	–0.014	–0.035	0.660	0.003	0.460	0.127	0.075	0.085
Median	0.002	0.012	0.000	–0.011	–0.040	0.669	–0.031	0.407	0.074	0.066	0.059
Standard Deviation	0.301	0.065	0.112	0.093	0.087	0.252	0.430	0.254	0.303	0.076	0.154
25th Percentile	0.000	0.004	–0.004	–0.047	–0.083	0.527	–0.275	0.275	0.001	0.036	0.010
75th Percentile	0.059	0.034	0.012	0.005	0.003	0.786	0.209	0.584	0.170	0.105	0.129
Panel C: Glamour Firms											
N	30,138	30,138	30,138	30,138	30,138	30,138	30,138	30,138	30,138	30,138	30,138
Mean	0.114	0.061	0.086	0.059	–0.010	0.782	0.352	0.389	0.340	0.050	0.136
Median	0.000	0.015	0.013	0.010	–0.019	0.735	0.174	0.346	0.141	0.035	0.068
Standard Deviation	0.277	0.107	0.200	0.141	0.107	0.429	0.773	0.256	0.663	0.113	0.248
25th Percentile	0.000	0.004	0.000	0.000	–0.068	0.545	–0.137	0.189	0.006	–0.013	0.009
75th Percentile	0.016	0.062	0.075	0.063	0.035	0.928	0.630	0.530	0.386	0.106	0.184

Notes: This exhibit reports summary statistics for all firm-year observations. The portfolios are formed on July 1 of year t *and are held until June 30 of year* t + 1*. Panel A shows summary statistics for the characteristics of all firms, and Panels B and C show the characteristics of low and high EM firms, respectively. DISTRESS is computed using the methodology described by Campbell, Hilscher, and Szilagyi (2008). OSCORE is computed using the methodology described by Ohlson (1980). NETISS is computed as the growth rate of the split-adjusted shares outstanding in the previous fiscal year. COMPISS is computed similarly to the method used by Daniel and Titman (2006). ACCRUAL is computed using the methodology of Sloan (1996). NOA is computed using the methodology of Hirshleifer et al. (2004). MOM is the cumulative returns from month −12 to month −2, as done by Fama and French (2008). GP is measured by gross profits scaled by AT, as by Novy-Marx (2013). AG is measured as the growth rate of the AT in the previous fiscal year, as by Cooper, Gulen, and Schill (2008). ROA is computed similarly to the method described by Piotroski and So (2012): Income before extraordinary items divided by AT INV is measured as the annual change in PPEGT plus the INVT scaled by the lagged book value of assets, as by Titman, Wei, and Xie (2004).*

(i.e., the 11 individual measures and the COMBO measure). The average monthly returns are significantly higher for all of the fundamental value proxies except for OSCORE, COMPISS, ACCRUAL, and ROA. These results are consistent with the hypothesis that the EM effect is likely explained by mispricing, not by enhanced systematic risk exposure.

We present a graphical depiction of the results from Exhibits 2 and 3, which plots the cumulative monthly returns from July 1, 1972 to December 31, 2015 for investments in (1) the risk-free asset, RF; (2) the hedge portfolio, EM_HML, constructed by taking a long position in the bottom quintile low EM value portfolio and a short position in the top quintile high EM glamour portfolio; (3) the low-mispricing portfolio, Low-Mispricing_COMBO, using the combination measure of fundamental value; and (4) the high-mispricing portfolio, High-Mispricing_COMBO, using the combination measure of fundamental value. The aggregate effects of the high-mispricing portfolio over time are dramatic, highlighting the behavioral nature of the EM effect.

In Exhibit 4, we present the details of the results of the last column in Exhibit 2, wherein we use the COMBO measure to identify the fundamental value of a specific portfolio. Specifically, Exhibit 4 shows the average monthly returns (Panel A) and alpha estimates (Panel B) for each of the 25 double-sorted portfolios using

EXHIBIT 2

Calendar-Time EM Portfolio Returns Conditional on Fundamental Value Proxies

	DISTRESS	OSCORE	NETISS	COMPISS	ACCRUAL	NOA	MOM	GP	AG	ROA	INV	COMBO
Panel A: EM Low-Mispricing												
Average Return	–0.13	0.63	0.37	0.40	0.26	0.03	–0.52	0.05	0.15	0.45	0.02	–0.06
Standard Deviation	7.13	5.68	4.67	4.90	5.69	5.96	8.52	5.18	5.06	5.62	5.02	5.65
CAPM Alpha	–0.10	**0.66**	0.36	0.26	0.34	0.13	–0.51	0.12	0.14	0.53	0.06	–0.14
	0.755	0.009	0.078	0.223	0.169	0.627	0.156	0.588	0.537	0.034	0.794	0.568
Three-Factor Alpha	**–0.61**	–0.03	0.05	0.01	–0.17	**–0.50**	**–1.10**	**–0.42**	0.11	–0.15	–0.09	**–0.61**
	0.042	0.879	0.787	0.965	0.433	0.023	0.002	0.028	0.634	0.428	0.703	0.007
Four-Factor Alpha	0.02	0.13	0.25	0.17	0.01	–0.24	0.22	–0.26	0.28	0.02	0.10	–0.17
	0.929	0.478	0.226	0.359	0.980	0.245	0.371	0.196	0.236	0.923	0.650	0.448
Panel B: EM High-Mispricing												
Average Return	1.30	0.52	0.98	0.83	0.70	0.92	1.99	0.91	1.02	0.61	0.82	1.31
Standard Deviation	6.33	5.88	5.22	5.47	5.82	5.00	7.84	4.77	6.10	6.55	5.54	5.76
CAPM Alpha	**1.54**	**0.72**	**1.15**	**1.09**	**0.89**	**1.07**	**2.31**	**0.93**	**1.29**	**0.86**	**1.04**	**1.60**
	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000
Three-Factor Alpha	**1.19**	**0.71**	**0.80**	**0.68**	0.37	**0.67**	**2.18**	**0.81**	**0.67**	**0.73**	**0.57**	**1.35**
	0.000	0.004	0.000	0.000	0.070	0.000	0.000	0.000	0.001	0.006	0.005	0.000
Four-Factor Alpha	**0.90**	**0.82**	**0.81**	**0.67**	**0.43**	**0.63**	**0.94**	**0.84**	**0.74**	**0.64**	**0.58**	**0.97**
	0.000	0.001	0.000	0.001	0.044	0.001	0.000	0.000	0.000	0.018	0.007	0.000
Paired *t*-Test (*p*-value)	0.001	0.777	0.047	0.179	0.216	0.009	0.000	0.005	0.013	0.680	0.015	0.000

Notes: This exhibit reports calendar-time portfolio regression alphas. The CAPM alpha is a risk-adjusted return equal to the intercept from a time-series regression of the long–short portfolios on the excess return on the value-weight market index (see Fama and French 1993). The three-factor alpha is a risk-adjusted return equal to the intercept from a time-series regression of the long–short portfolios on the excess return on the value-weight market index, the return on the size (SMB) factor, and the return on the value (HML) factor (see Fama and French 1993). The four-factor alpha is a risk-adjusted return equal to the intercept from a time-series regression of long–short portfolios on the excess return on the value-weight market index, the return SMB, the return on HML, and the return on a prior-year return momentum (MOM) factor (see Carhart 1997). Average alphas are in monthly percentage, p*-values are shown below the coefficient estimates, and 5% statistical significance is indicated in bold. Portfolios for each strategy are rebalanced monthly. Panel A presents a low-mispricing portfolio that is long value and short glamour firms with the lowest investor expectation errors. Panel B presents a high-mispricing portfolio that is long value and short glamour firms with the highest investor expectation errors. The final row is a* p*-value from a paired* t*-test comparing the high-mispricing portfolio monthly returns (Panel B) and the low-mispricing portfolio monthly returns (Panel A). The time period under analysis is from July 1, 1972 to December 31, 2015. Regression* p*-values use robust standard errors as computed by Davidson and MacKinnon (1993, p. 553).*

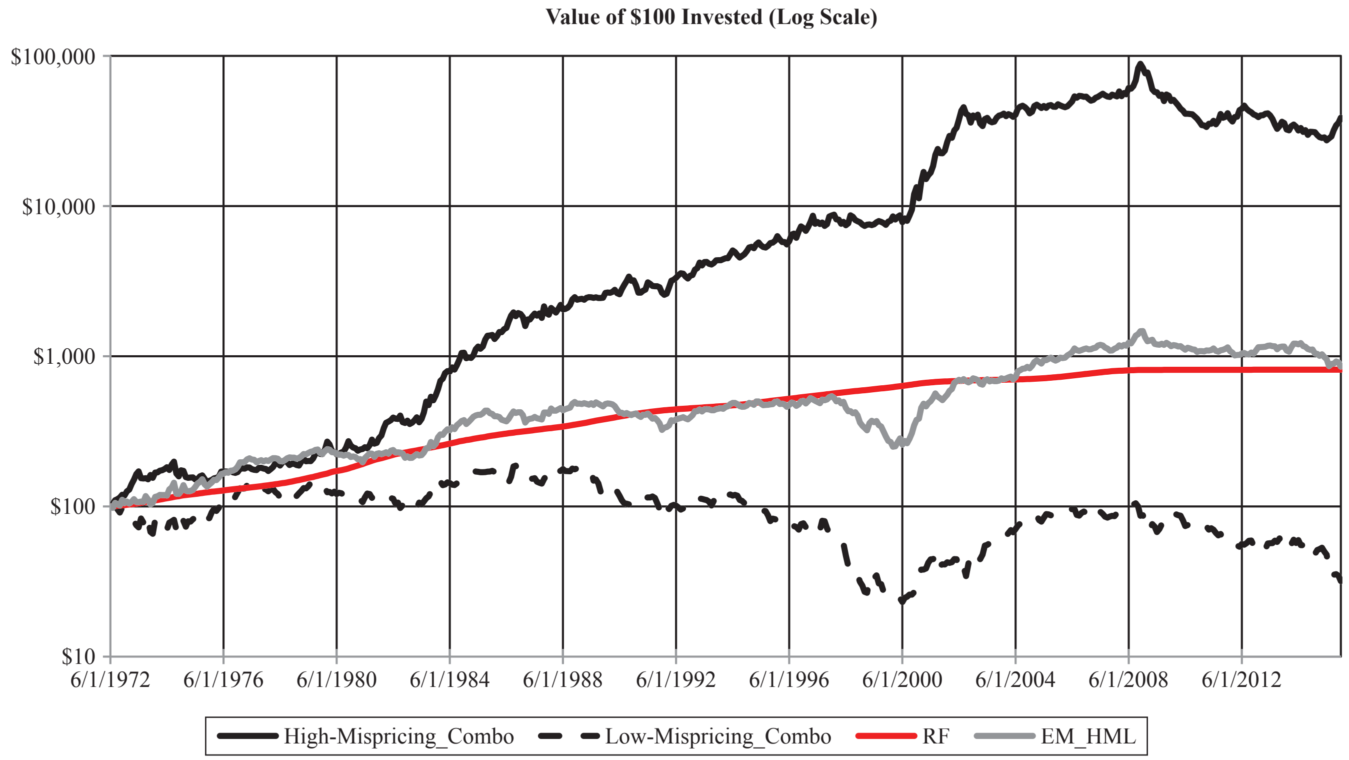
EM and COMBO. We include the results in Exhibit 4 to give the reader a better sense of the source of the returns and alphas reported in Exhibit 2. The returns and alpha estimates increase across the fundamental value quintiles. Furthermore, consistent with the EM anomaly, alphas increase across the EM quintiles. Holding the EM quintile constant, alphas generally increase across fundamental value quintiles.

Realized Expectation Errors to EM Conditional on Fundamental Value Proxies

To assess whether our fundamental value proxies are capturing mispricing effects within EM-sorted portfolios, we analyze revisions to market expectations subsequent to portfolio formation. Following Piotroski and So (2012), we examine earnings announcement period returns, analyst forecast errors, and analyst forecast revisions to see whether these variables vary when expectations embedded in price (EM portfolios) are consistent with expectations from firm fundamentals. If mispricing is driving the EM effect, we should see higher earnings announcement returns when expectation errors are high (value firms with high fundamental value) than when they are low (glamour firms with low fundamental value). Furthermore, we should see a positive spread in forecast errors and forecast revisions across high- and low-expectation-error firms, controlling for valuation (i.e., EM quintile).

Exhibit 3

Cumulative Gains from Investment for High-Mispricing and Low-Mispricing EM Portfolios



Notes: Plotted are the cumulative returns for four assets: (1) the risk-free asset, RF; (2) the hedge portfolio, EM_HML, constructed by taking a long position in the bottom quintile low EM value portfolio and a short position in the top quintile high EM glamour portfolio; (3) the low-mispricing portfolio, Low-Mispricing_Combo, using the combination measure of fundamental value; and (4) the high-mispricing portfolio, High-Mispricing_Combo, using the combination measure of fundamental value. The sample period is from July 1, 1972 to December 31, 2015.

The logic behind examining earnings announcement returns subsequent to portfolio formation is demonstrated by La Porta, Shleifer, and Vishny (1997). They showed that glamour firms experience negative earnings announcement returns following portfolio formation, consistent with the market revising downward its overly optimistic expectations about glamour firms when they announce earnings; the opposite is true for value firms.

We calculate the three-day market-adjusted earnings announcement return during the fourth quarter for a given firm in the year following portfolio formation.[6] We then focus on the average earnings announcement returns in high-mispricing and low-mispricing portfolios. The mispricing hypothesis predicts that subsequent earnings announcement returns will be larger for the high-mispricing portfolio relative to the low-mispricing portfolio; the risk-based hypothesis predicts that they will be similar. Exhibit 5, Panel A presents the results. The results show that low EM value firms have higher earnings announcement returns than high EM glamour firms across all fundamental value quintiles. In addition, in every EM quintile, earnings announcement returns are larger for firms in quintile five of our composite fundamental value measure than for firms in quintile one, suggesting that the market is surprised by subsequent earnings announcements for firms with high fundamental value. Most importantly, the average announcement return for the value portfolio with high expectation error (i.e., cheap with high fundamental value) is 0.65%, which is larger than the −0.22% average announcement return for the glamour

[6] Results are similar if we use all the earnings announcement returns for all firm quarters in the year following portfolio formation.

EXHIBIT 4
Average Monthly EM Returns Conditional on Fundamental Value Proxy

Panel A: EM Average Monthly Returns

	EM Quintiles						
FV Quintiles	**Glamour (1)**	**2**	**3**	**4**	**Value (5)**	**V-G Diff.**	***p*-Value**
Low (1)	–0.01	0.46	0.62	0.97	0.97	**0.98**	0.025
2	0.65	0.67	0.99	0.97	1.17	0.53	0.182
3	0.67	0.81	0.93	1.15	1.48	**0.81**	0.025
4	1.09	0.92	0.99	1.18	1.32	0.23	0.542
High (5)	1.03	0.96	1.13	1.15	1.30	0.27	0.375
High-Low	**1.04**	0.50	0.51	0.19	0.33		
(*p*-value)	0.008	0.121	0.120	0.566	0.361		
Low-Mispricing						–0.06	0.880
High-Mispricing						**1.31**	0.000

Panel B: Alpha Estimates

	EM Quintiles				
	Glamour (1)	**2**	**3**	**4**	**Value (5)**
CAPM Error Quintiles					
Low (1)	–1.11	–0.51	–0.37	–0.01	–0.05
2	–0.43	–0.27	0.14	0.06	0.23
3	–0.35	–0.13	0.06	0.24	0.59
4	0.05	0.02	0.11	0.34	0.44
High (5)	0.09	0.08	0.28	0.33	0.49
Three-Factor Error Quintiles					
Low (1)	–1.01	–0.61	–0.56	–0.28	–0.30
2	–0.25	–0.32	–0.05	–0.18	–0.04
3	–0.11	–0.15	–0.09	0.01	0.39
4	0.30	0.03	–0.01	0.16	0.20
High (5)	0.31	0.15	0.22	0.19	0.34
Four-Factor Error Quintiles					
Low (1)	–0.70	–0.35	–0.20	0.00	0.05
2	–0.10	–0.20	0.06	0.02	0.16
3	–0.09	–0.10	–0.09	0.09	0.52
4	0.23	0.01	–0.03	0.22	0.27
High (5)	0.22	0.12	0.15	0.17	0.27

Notes: This exhibit reports average monthly returns (Panel A) and calendar-time portfolio regression alphas (Panel B) for portfolios sorted on the composite fundamental value proxy (FV Quintiles), conditional on enterprise multiples (EM Quintiles). The CAPM alpha is a risk-adjusted return equal to the intercept from a time-series regression of the long–short portfolios on the excess return on the value-weight market index (see Fama and French 1996). The three-factor alpha is a risk-adjusted return equal to the intercept from a time-series regression of the long–short portfolios on the excess return on the value-weight market index, the return on the size (SMB) factor, and the return on the value (HML) factor (see Fama and French 1996). The four-factor alpha is a risk-adjusted return equal to the intercept from a time-series regression of long–short portfolios on the excess return on the value-weight market index, the return on the size (SMB) factor, the return on the value (HML) factor, and the return on a prior-year return momentum (MOM) factor (see Carhart 1997). Average alphas are in monthly percent and 5% statistical significance is indicated in bold. Portfolios for each strategy are rebalanced monthly. Panels A highlights average monthly returns across the portfolio sorts. Panel B shows the portfolio alpha estimates for various factor models. The time period under analysis is from July 1, 1972 to December 31, 2015. Difference tests are paired t*-tests for difference. Alpha estimate* p*-values use robust standard errors as computed by Davidson and MacKinnon (1993, p. 553).*

EXHIBIT 5

Realized Expectation Errors to Fundamental Value Proxies Conditional on EM

FV Quintiles	EM Quintiles: Glamour (1)	2	3	4	Value (5)	V-G Diff.	*p*-Value
Panel A: EM Earnings Announcement Returns							
Low (1)	–0.22	0.01	0.30	0.43	0.32	**0.54**	0.022
2	0.01	0.57	0.47	0.56	0.53	**0.52**	0.012
3	0.36	0.50	0.56	0.63	0.67	0.31	0.125
4	0.44	0.51	0.74	0.64	0.80	0.36	0.051
High (5)	0.50	0.54	0.74	0.60	0.65	0.15	0.360
High-Low	**0.72**	**0.53**	**0.44**	0.18	0.33		
(*p*-value)	0.000	0.003	0.020	0.323	0.077		
Low-Mispricing V/G						–0.18	0.368
High-Mispricing V/G						**0.87**	0.000
Panel B: EM Forecast Errors							
Low (1)	–0.031	–0.021	–0.018	–0.021	–0.030	0.001	0.800
2	–0.018	–0.014	–0.010	–0.012	–0.020	–0.002	0.317
3	–0.013	–0.014	–0.010	–0.009	–0.015	–0.001	0.483
4	–0.010	–0.005	–0.005	–0.004	–0.013	–0.002	0.227
High (5)	–0.003	–0.002	–0.003	–0.003	–0.008	**–0.005**	0.002
High-Low	**0.028**	**0.019**	**0.015**	**0.018**	**0.022**		
(*p*-value)	0.00	0.00	0.00	0.00	0.00		
Low-Mispricing V/G						**–0.027**	0.000
High-Mispricing V/G						**0.023**	0.000
Panel C: EM Forecast Revisions							
Low (1)	–0.021	–0.013	–0.012	–0.013	–0.020	0.001	0.792
2	–0.012	–0.010	–0.007	–0.010	–0.014	**–0.003**	0.029
3	–0.010	–0.007	–0.007	–0.007	–0.011	–0.001	0.291
4	–0.007	–0.004	–0.005	–0.005	–0.009	–0.002	0.138
High (5)	–0.002	–0.001	–0.002	–0.003	–0.006	**–0.004**	0.000
High-Low	**0.019**	**0.012**	**0.010**	**0.010**	**0.013**		
(*p*-value)	0.000	0.000	0.000	0.000	0.000		
Low-Mispricing V/G						**–0.018**	0.000
High-Mispricing V/G						**0.015**	0.000

Notes: This exhibit presents earnings announcement returns (fourth quarter), consensus analyst forecast errors (FE), and revisions (REV) for portfolios sorted on the composite fundamental value proxy (FV quintiles), conditional on enterprise multiples (EM quintiles). Analyst forecast errors and revisions are calculated six months after the preceding fiscal year's end. FE is defined as (Actual EPS – Consensus forecast)/(Total assets per share), and forecast REV is defined as the final consensus estimate minus the consensus at portfolio formation scaled by total assets per share. The low-mispricing V/G strategy consists of a long position in value firms with low fundamental value and a short position in glamour firms with high fundamental value. The high-mispricing V/G strategy consists of a long position in value firms with high fundamental value and a short position in glamour firms with low fundamental value. The time period under analysis is from July 1, 1972 to December 31, 2015. Difference tests are paired t*-tests for difference. 5% statistical significance is indicated in bold.*

portfolio with high expectation error (i.e., expensive with low fundamental value). The difference in these two returns is the announcement return of 0.87% in the high-mispricing portfolio, which is significant at the 1% level. Alternatively, the returns to firms in each of the two portfolios with low expectation errors are similar: Value firms with low fundamental value have subsequent announcement returns of 0.32%, whereas the average earnings announcement return of glamour firms with high fundamental value is 0.50%.

We gather data from IBES to calculate analyst forecast errors and revisions. Not all firms are covered by analysts, so we lose a significant portion of the sample in our analysis of analyst forecast properties. We calculate analyst forecast errors as the actual annual earnings per share (EPS) at the end of the year less the consensus EPS in the month preceding portfolio formation and scale the difference by AT per share. We calculate analyst forecast revisions as the consensus forecast immediately preceding the annual earnings announcement date minus the consensus forecast from the month prior to portfolio formation, scaled by AT per share. Exhibit 5, Panel B (Panel C) shows the results for forecast errors (revisions).

Negative forecast errors indicate that analysts are optimistic with respect to realized earnings, whereas positive forecast errors indicate that analysts are pessimistic. On average, analysts are more optimistic with respect to firms with low fundamental value than firms with high fundamental value. More importantly, analysts following the value firms with the highest fundamental value are much less optimistic than analysts following glamour firms with low fundamental value, which is consistent with mispricing driving the returns in the high-mispricing portfolio. Specifically, the difference in average forecast errors for these two groups is 0.023, which is significant at the 1% level.

Turning to Panel C, we observe similar results for forecast revisions. Negative forecast revisions suggest that analysts revise their earnings expectations downward over time; the more negative the number, the more drastic the downward revision. As with forecast errors, the average firm in our sample experiences negative forecast revisions. In the high-mispricing portfolio, we observe small negative forecast revisions for value firms with high fundamental value relative to glamour firms with low fundamental value. Again, this suggests that analysts do not have to revise their forecasts downward as much for value firms that are not expected to perform well but which have high fundamental value. The combined results in Exhibit 5 provide additional evidence that the positive (negative) returns of low (high) EM firms are consistent with market mispricing.

EM Portfolio Performance, Conditional on Investor Sentiment

We conduct an additional test to examine the mispricing hypothesis. Following Stambaugh, Yu, and Yuan (2012), we introduce a measure of investor sentiment to determine whether the returns to our high-mispricing portfolio are stronger in periods of high market sentiment. The mispricing hypothesis predicts that returns to a high-mispricing portfolio will be stronger in periods when expectation errors are predicted to be the most extreme (i.e., during high market sentiment). We use two measures of market sentiment. The first is the measure developed by Baker and Wurgler (2006) as the first principal component of six proxies of investor sentiment. The second is a measure developed by Huang et al. (2014) that builds and improves upon the measure proposed by Baker and Wurgler (2006). Specifically, the Huang et al. (2014) measure uses a partial least squares method to separate relevant information from the proxies from noise. Exhibit 6 shows the average monthly returns to the low-mispricing and high-mispricing EM portfolios across periods of low, medium, and high investor sentiment. In support of the mispricing hypothesis, the returns to the high-mispricing EM strategy are significantly higher during periods of high investor sentiment relative to times of low investor sentiment; the same pattern is not observed for the low-mispricing EM portfolio. For the simple value/glamour EM-sorted portfolios (V-G in Exhibit 6), returns are marginally higher during periods of high investor sentiment.

Robustness Tests

We summarize three important tests that we perform to assess the robustness of our results. The first test examines results across two different time periods in our sample. In our second test, we examine whether the results are robust to the January effect documented in prior literature. Finally, we introduce three other multifactor models to ensure that the EM premium cannot be explained by other controls for risk. These models include the q-theory factor model (Hou, Xue, and Zhang 2017), the Fama–French five-factor model (Fama and French 2015), and the AQR six-factor model (Asness 2014).

For the first test, we split our sample into two distinct time periods. The first time period extends from July 1972 to December 1993. The latter period begins in January 1994 and ends in December 2015. We examine the cross-sectional regressions on EM portfolios in both subsamples. The results in both time periods are quantitatively similar to those documented in Exhibit 2, which represents the full sample period. This analysis minimizes the worry that the results are driven by sample selection.

EXHIBIT 6
EM Long–Short Portfolio Returns Conditional on Level of Investor Sentiment

	Baker and Wurgler (2006)			Huang et al. (2014)		
Sentiment	**V-G**	**Low-Mispricing**	**High-Mispricing**	**V-G**	**Low-Mispricing**	**High-Mispricing**
Low	0.495	0.675	0.333	0.457	0.043	0.681
Medium	0.005	–0.694	0.598	0.253	0.124	0.591
High	1.090	0.342	3.188	0.799	–0.106	2.833
High-Low	0.595	–0.333	**2.855**	0.342	–0.149	**2.152**
p-Value	0.184	0.615	0.000	0.445	0.811	0.002

Notes: This exhibit presents average monthly returns, conditional on the level of investor sentiment in the market, from July 1, 1972 to December 31, 2014. Investor sentiment reflects the index used by Baker and Wurgler (2006) and the index used in Huang et al. (2014). Both sentiment indexes are orthogonalized to macro factors. Investor sentiment is measured in the month preceding portfolio formation. The low-mispricing portfolio is long value and short glamour firms with the lowest investor expectation errors, the high-mispricing portfolio is long value and short glamour firms with the highest investor expectation errors, and the V-G portfolio is long value and short glamour firms, independent of an expectation error proxy. The final row is a p*-value from a paired* t*-test comparing the high portfolio monthly returns and the low monthly returns, and 5% statistical significance is indicated in bold.*

For the second robustness test, we examine the January effect. Several researchers have documented that some anomalies' returns are concentrated in the month of January. To ensure that the EM effect is not simply a manifestation of this seasonality in returns, we repeat our analysis after removing all January returns. The results are robust to removal of these observations, suggesting that the January effect does not drive our core findings.

In our third robustness tests, we redo our results using three more recently developed factor models: (1) the q-factor model, described by Hou, Xue, and Zhang (2017), which includes Mkt-Rf, ME (size), I/A (investment-to-assets) and ROE (return on equity); (2) the Fama–French five-factor model (Fama and French 2015), which includes the Fama–French three factors and profitability and investment factors; and (3) a six-factor model described by Asness (2014), which includes Mkt-Rf, SMB, HML-DEV (an adjusted value factor), RMW (robust minus weak), CMA (conservative minus aggressive), and UMD (winners minus losers). Our results are robust to these factor models.

The summary of our additional tests is that our core results are robust to sample period, the January calendar effect, and different asset pricing models.

Limits of Arbitrage and the EM Effect

The evidence presented suggests that the EM effect is driven by mispricing. The sample period tests from the previous section show that the mispricing estimates are essentially the same across the early and latter halves of the sample. Given the evidence for mispricing, however, why have market participants not fully exploited the EM effect? The theory of limits to arbitrage suggests that if an arbitrageur's cost to exploit a mispricing is too high, then the mispricing may not be a transitory phenomenon (e.g., Shleifer and Vishny 1997). For example, Brav, Heaton, and Li (2010) found that limits to arbitrage can explain overvaluation (glamour stocks) but not undervaluation (value stocks). One simple way to test how limits of arbitrage affect the returns of the high-mispricing portfolio is to examine the mispricing generated by the long and short legs independently. We assume that arbitrage costs are generally higher for short portfolios than they are for long portfolios. The limits of arbitrage hypothesis predicts that the abnormal expected returns associated with the high-mispricing portfolio will be driven by the short leg of the portfolio where arbitrage costs are highest.

Exhibit 7 shows the performance results for the long and short legs of the high-mispricing portfolios. An analysis of raw returns suggests that the absolute performance of the high-mispricing EM portfolio is driven by the long leg of the portfolio. However, the risk-adjusted results tell a different story: the long book of the high-mispricing portfolio generates, on average, 38% of the three-factor alpha (Exhibit 7, Panel A), whereas 62% of the three-factor alpha comes from the short side of the high-mispricing portfolio (Exhibit 7, Panel B).

EXHIBIT 7

Limits of Arbitrage: Long and Short Legs of High-Mispricing EM Portfolio

	DISTRESS	OSCORE	NETISS	COMPISS	ACCRUAL	NOA	MOM	GP	AG	ROA	INV	COMBO
Panel A: EM High-Mispricing Long Leg												
Average Return	1.52	1.29	1.45	1.36	1.17	1.35	1.52	1.39	1.67	1.33	1.50	1.30
Standard Deviation	5.54	5.19	5.24	4.80	5.60	5.60	5.42	5.99	5.39	5.81	5.33	4.66
CAPM Alpha	**0.66**	**0.41**	**0.59**	**0.55**	0.26	**0.43**	**0.64**	**0.46**	**0.79**	**0.42**	**0.63**	**0.49**
	0.000	0.003	0.000	0.000	0.090	0.004	0.000	0.006	0.000	0.011	0.000	0.000
Three-Factor Alpha	**0.39**	**0.30**	**0.35**	**0.28**	0.04	0.20	**0.48**	0.25	**0.48**	0.29	**0.36**	**0.34**
	0.016	0.029	0.014	0.024	0.771	0.154	0.002	0.101	0.001	0.065	0.007	0.006
Four-Factor Alpha	0.33	**0.37**	**0.41**	**0.30**	0.12	**0.29**	0.14	**0.38**	**0.59**	0.31	**0.40**	**0.27**
	0.053	0.010	0.005	0.020	0.433	0.045	0.332	0.013	0.000	0.060	0.005	0.037
Panel B: EM High-Mispricing Short Leg												
Average Return	0.22	0.77	0.47	0.53	0.47	0.43	–0.47	0.48	0.65	0.72	0.68	–0.01
Standard Deviation	7.56	7.60	6.41	6.64	7.27	6.71	8.92	5.75	7.58	8.36	6.83	7.20
CAPM Alpha	**–0.89**	–0.31	**–0.57**	**–0.54**	**–0.64**	**–0.64**	**–1.67**	**–0.47**	**–0.50**	**–0.44**	**–0.41**	**–1.11**
	0.000	0.139	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.046	0.005	0.000
Three-Factor Alpha	**–0.80**	**–0.41**	**–0.46**	**–0.40**	**–0.33**	**–0.47**	**–1.70**	**–0.56**	–0.19	**–0.44**	–0.20	**–1.01**
	0.000	0.045	0.001	0.003	0.020	0.001	0.000	0.000	0.200	0.035	0.140	0.000
Four-Factor Alpha	**–0.58**	**–0.45**	**–0.40**	**–0.37**	**–0.31**	**–0.35**	**–0.81**	**–0.46**	–0.15	**–0.33**	–0.19	**–0.70**
	0.003	0.033	0.004	0.006	0.037	0.013	0.000	0.002	0.303	0.135	0.199	0.000
Paired *t*-test (*p*-value)	0.002	0.194	0.007	0.021	0.080	0.017	0.000	0.013	0.012	0.172	0.031	0.001
Short CAPM Alpha %	0.58	0.43	0.49	0.50	0.71	0.60	0.72	0.51	0.39	0.51	0.40	0.70
Short Three-Factor Alpha %	0.67	0.58	0.57	0.59	0.88	0.70	0.78	0.69	0.28	0.60	0.36	0.75
Short Four-Factor Alpha %	0.64	0.55	0.49	0.55	0.73	0.55	0.85	0.55	0.20	0.51	0.32	0.72

Notes: This exhibit reports calendar-time portfolio regression alphas. The CAPM alpha is a risk-adjusted return equal to the intercept from a time-series regression of the long–short portfolios on the excess return on the value-weight market index (see Fama and French 1996). The three-factor alpha is a risk-adjusted return equal to the intercept from a time-series regression of the long–short portfolios on the excess return on the value-weight market index, the return on the size (SMB) factor, and the return on the value (HML) factor (see Fama and French 1996). The four-factor alpha is a risk-adjusted return equal to the intercept from a time-series regression of long–short portfolios on the excess return on the value-weight market index, the return on SMB, the return on HML factor, and the return on a prior-year return momentum (MOM) factor (see Carhart 1997). Average alphas are in monthly percentage, p*-values are shown below the coefficient estimates, and 5% statistical significance is indicated in bold. Portfolios for each strategy are rebalanced monthly. Panel A represents the long leg of the high-mispricing portfolio that is long value with the highest investor expectation errors. Panel B represents the short leg of the high-mispricing portfolio that is short glamour with the highest investor expectation errors. We display the* p*-value from a paired* t*-test comparing the high-mispricing portfolio monthly returns (Panel B) and the low-mispricing portfolio monthly returns (Panel A) below Panel B and the percentage of long–short alpha from the short book. The time period under analysis is from July 1, 1972 to December 31, 2015. Regression* p*-values use robust standard errors as computed by Davidson and MacKinnon (1993, p. 553).*

The empirical results suggest that the absolute pricing is concentrated in the long leg of the high-mispricing EM portfolio; however, on a risk-adjusted basis, the mispricing is concentrated in the short portfolio of glamour firms for the majority of the high-mispricing portfolio constructs. Returns associated with short-selling strategies are often difficult to obtain in practice. For instance, Beaver, McNichols, and Price (2016) found that short selling entails significant costs that affect trading strategy profitability. To the extent that managing short positions is costly, these results suggest that the mispricing associated with the high-mispricing EM portfolio is difficult to profitably exploit.

We should also note that the potential mispricing identified does not represent an easy profit opportunity on either the long or the short leg of the portfolio. The average monthly returns, although favorable, are highly volatile (see Exhibit 2). To the extent one believes there are principle–agent conflicts between asset owners and fund managers (discussed by Shleifer and Vishny 1997), one would expect that strategies seeking to exploit longer-term profit opportunities, which includes the EM value anomaly discussed in this article, could continue to exist in the future. Observing the EM effect in theory may be a lot easier than exploiting the opportunity in practice

CONCLUSION

The evidence supports the hypothesis that the excess returns associated with EM-sorted portfolios is driven by mispricing and not by increased systematic risk exposure. The EM effect is stronger among portfolios that sort on fundamental value proxies to identify portfolios with predictable investor expectation errors. Moreover, we document that earnings announcement returns, forecast errors, and forecast revisions all suggest that mispricing likely drives the EM effect. Finally, we show that the EM effect is much stronger during times of strong market sentiment, which also supports the mispricing-based hypothesis.

ACKNOWLEDGMENTS

We would like to thank David Foulke, Christian Walkshäusl, Eric So, and Andrew Miller for comments and suggestions.

REFERENCES

Asness, C. "Our Model Goes to Six and Saves Value from Redundancy Along the Way." AQR.com, 2014, https://www.aqr.com/cliffs-perspective/our-model-goes-to-six-and-saves-value-from-redundancy-along-the-way.

Baker, M., and J. Wurgler. 2006. "Investor Sentiment and the Cross-Section of Stock Returns." *The Journal of Finance* 61 (4): 1645–1680.

Beaver, W., M. McNichols, and R. Price. 2007. "Delisting Returns and Their Effect on Accounting-Based Market Anomalies." *Journal of Accounting and Economics* 43 (2–3): 341–368.

——. 2016. "The Cost and Benefits of Long–Short Investing: A Perspective on the Market Efficiency Literature." *Journal of Accounting Literature* 37 (2016): 1–18.

Brav, A., J. B. Heaton, and S. Li. 2010. "The Limits of the Limits of Arbitrage." *The Review of Finance* 14 (1): 157–187.

Campbell, J., J. Hilscher, and J. Szilagyi. 2008. "In Search of Distress Risk." *The Journal of Finance* 63 (6): 2899–2939.

Campbell, J., C. Polk, and T. Vuolteenaho. 2010. "Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns." *Review of Financial Studies* 23 (1): 305–344.

Carhart, M. 1997. "On Persistence in Mutual Fund Performance." *The Journal of Finance* 52 (1): 57–82.

Chen, L., R. Novy-Marx, and L. Zhang. 2010. "An Alternative Three-Factor Model." Working paper.

Cooper, M., H. Gulen, and M. J. Schill. 2008. "Asset Growth and the Cross-Section of Stock Returns." *The Journal of Finance* 63 (4): 1609–1652.

Daniel, K., and S. Titman. 2006. "Market Reactions to Tangible and Intangible Information." *The Journal of Finance* 61 (4): 1605–1643.

Davidson, R., and J. MacKinnon. *Estimation and Inference in Econometrics*. New York: Oxford University Press, 1993.

Fama, E. 1998. "Market Efficiency, Long-Term Returns, and Behavioral Finance." *Journal of Financial Economics* 49 (3): 283–306.

Fama, E., and K. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33 (1): 3–56.

——. 1996. "Multifactor Explanations of Asset Pricing Anomalies." *The Journal of Finance* 51 (1): 55–84.

——. 2006. "Profitability, Investment, and Average Returns." *Journal of Financial Economics* 82 (3): 491–518.

——. 2008. "Dissecting Anomalies." *The Journal of Finance* 63 (4): 1653–1678.

——. 2015. "A Five-Factor Asset Pricing Model." *Journal of Financial Economics* 116 (1): 1–22.

Gray, W., and J. Vogel. 2012. "Analyzing Valuation Measures: A Performance Horse Race Over the Past 40 Years." *The Journal of Portfolio Management* 39 (1): 112–121.

Gray, W., and T. E. Carlisle. *Quantitative Value: A Practitioner's Guide to Automating Intelligent Investment and Eliminating Behavioral Errors*. New Jersey: John Wiley and Sons, 2012.

Hirshleifer, D., K. Hou, S. Teoh, and Y. Zhang. 2004. "Do Investors Overvalue Firms with Bloated Balance Sheets." *Journal of Accounting and Economics* 38 (4–5): 297–331.

Hou, K. C. Xue, and L. Zhang. "A Comparison of New Factor Models." Working paper, 2017.

Huang, D., F. Jiang, J. Tu, and G. Zhou. 2014. "Investor Sentiment Aligned: A Powerful Predictor of Stock Returns." *The Review of Financial Studies* 28 (3): 791–837.

Jegadeesh, N., and S. Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *The Journal of Finance* 48 (1): 65–91.

Lakonishok, J., A. Shleifer, and R. Vishny. 1994. "Contrarian Investment, Extrapolation, and Risk." *The Journal of Finance* 49 (5): 1541–1578.

La Porta, R., A. Shleifer, and R. Vishny. 1997. "Good News for Value Stocks: Further Evidence on Market Efficiency." *The Journal of Finance* 52 (2): 859–874.

Loughran, T., and J. Ritter. 1995. "The New Issues Puzzle." *The Journal of Finance* 50 (1): 23–51.

Loughran, T., and J. Wellman. 2011. "New Evidence on the Relation between the Enterprise Multiple and Average Stock Returns." *Journal of Financial and Quantitative Analysis* 46 (6): 1629–1650.

Mitchell, M., and E. Stafford. 2000. "Managerial Decisions and Long-Term Stock Price Performance." *The Journal of Business* 73 (3): 287–329.

Novy-Marx, R. 2013. "The Other Side of Value: Good Growth and the Gross Profitability Premium." *Journal of Financial Economics* 108 (1): 1–28.

Ohlson, J. 1980. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." *Journal of Accounting Research* 18 (1): 109–131.

Piotroski, J. 2000. "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers." *Journal of Accounting Research* 38: 1–41.

Piotroski, J., and E. So. 2012. "Identifying Expectation Errors in Value/Glamour Strategies: A Fundamental Analysis Approach." *The Review of Financial Studies* 25 (9): 2841–2875.

Ritter, J. 1991. "The Long-Run Performance of Initial Public Offerings." *The Journal of Finance* 46 (1): 3–27.

Santos, T., and P. Veronesi. 2010. "Habit Formation, the Cross-Section of Stock Returns and the Cash-Flow Risk Puzzle." *Journal of Financial Economics* 98 (2): 385–413.

Shleifer, A., and R. Vishny. 1997. "The Limits of Arbitrage." *The Journal of Finance* 52 (1): 35–55.

Sloan, R. 1996. "Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings?" *The Accounting Review* 71 (3): 289–315.

Stambaugh, R., J. Yu, and Y. Yuan. 2012. "The Short of It: Investor Sentiment and Anomalies." *Journal of Financial Economics* 104 (2): 288–302.

Titman, S., K. Wei, and F. Xie. 2004. "Capital Investments and Stock Returns." *Journal of Financial and Quantitative Analysis* 39 (4): 677–700.

Walkshäusl, C., and S. Lobe. 2015. "The Enterprise Multiple Investment Strategy: International Evidence." *Journal of Financial and Quantitative Analysis* 50 (4): 781–800.

Xing, Y. 2008. "Interpreting the Value Effect through the Q-Theory: An Empirical Investigation." *The Review of Financial Studies* 21 (4): 1767–1795.

ADDITIONAL READING

Enhancing the Investment Performance of Yield-Based Strategies
Wesley R. Gray and Jack Vogel
The Journal of Investing
https://joi.pm-research.com/content/23/2/44

ABSTRACT: *High-dividend-yield stocks do not reliably earn above-average risk-adjusted returns. More complete measures of shareholder yield, which account for net share repurchases, perform better. This article explores the use of net-debt paydown as a way to further enhance shareholder yield. The addition of net-debt paydown enhances risk-adjusted returns and creates a shareholder yield metric that is more robust across time and to the inclusion or exclusion of financials.*

Analyzing Valuation Measures: *A Performance Horse Race over the Past 40 Years*

Wesley R. Gray and Jack Vogel

The Journal of Portfolio Management
https://jpm.pm-research.com/content/39/1/112

ABSTRACT: *In this article, the authors compare the investment performance of portfolios sorted on different valuation measures. They find that EBITDA/TEV has been the best performing metric, historically, and outperforms many investor favorites, such as price-to-earnings ratio, free cash flow to total enterprise value, and book-to-market ratio. The authors also explore the investment potential of long-term valuation ratios, which replace one-year earnings with an average of long-term earnings. They find that in contrast to prior empirical work, long-term ratios add little investment value over standard one-year valuation metrics.*

Do You Know What's in Your Benchmark?

Steven Crawford, James Hansen, and Richard Price

The Journal of Portfolio Management
https://jpm.pm-research.com/content/39/3/136

ABSTRACT: *The authors identify several problematic assumptions underlying the benchmark return methodology used by the Center for Research in Security Prices (CRSP), which practitioners and academics would be unlikely to know or mimic. In particular, CRSP includes non-common stock securities that most researchers exclude. CRSP does not follow a typical buy-and-hold methodology, and it excludes delisting returns. The authors discuss these issues and show how they can affect results in a number of research settings. The commonly used value-weighted, size-based benchmark returns, as well as all equally weighted daily benchmark returns, are particularly problematic.*