Journal of Financial Economics 117 (2015) 470–488

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Incremental variables and the investment opportunity set ☆

Eugene F. Fama [a], Kenneth R. French [b,*]

[a] Booth School of Business, University of Chicago, Chicago, IL 60637, USA
[b] Amos Tuck School of Business, Dartmouth College, Hanover, NH 03750, USA

ARTICLE INFO

Article history:
Received 24 September 2013
Received in revised form
13 October 2014
Accepted 24 December 2014
Available online 13 May 2015

JEL classification:
G11
G12

Keywords:
Incremental variables
Investment opportunity set
Portfolio returns
Variable attenuation

ABSTRACT

Variables with strong marginal explanatory power in cross-section asset pricing regressions typically show less power to produce increments to average portfolio returns, for two reasons. (1) Adding an explanatory variable can attenuate the slopes in a regression. (2) Adding a variable with marginal explanatory power always attenuates the values of other explanatory variables in the extremes of a regression's fitted values. Without a restriction on portfolio weights, the maximum Sharpe ratios in the GRS statistic of Gibbons, Ross, and Shanken (1989) provide little information about an incremental variable's impact on the portfolio opportunity set.

1. Introduction

Many asset pricing papers identify individual variables (e.g., size, the book-to-market ratio, momentum, accruals, net share issues – the list is long) that are related to average returns. A variable's importance is often judged by the spread it produces in average returns. The emphasis is typically on the extreme quintiles or deciles from a sort on the variable, often with eye-popping results. Given the many patterns already identified in the cross section of average returns, however, a new variable's univariate average return spread almost certainly overstates its incremental contribution to the average return spread for portfolios formed using multiple forecasting variables.

Our goal is to explain a variable's incremental contribution to the average return spread from a multivariate forecast. We take the perspective of a researcher who uses estimates of expected returns from Fama-MacBeth (FM, 1973) cross-section regressions to sort stocks into portfolios. A variable's incremental impact on the spread in average returns from a multivariate regression is usually smaller than the spread from a univariate sort for two related reasons. First, a new explanatory variable often attenuates the slopes of variables already in the regression. Slope attenuation occurs in a bivariate regression, for example, when the explanatory variables are positively correlated and their slopes have the same sign or when they are negatively correlated and their slopes have opposite signs. The result is a reduction in the incremental variable's impact on the expected return spread.

The second driver is variable attenuation. Adding a variable with marginal explanatory power always shrinks the values of other explanatory variables in the extremes of a regression's fitted values. As a result, getting the expected return benefits of an additional variable almost always involves losing some of the gains from variables

☆ Eugene F. Fama and Kenneth R. French are stockholders in, board members of, and consultants to Dimensional Fund Advisors. We thank participants at the 2012 Finance Down Under Conference and the Norwegian School of Economics and Business Administration Finance Seminar, and an anonymous referee for helpful comments.

* Corresponding author.
E-mail address: Kenneth.R.French@tuck.dartmouth.edu (K.R. French).

http://dx.doi.org/10.1016/j.jfineco.2015.05.001
0304-405X/

already in the mix. The intuition is straightforward. Sorting on the fitted values from a one-variable regression is equivalent to sorting on the variable itself, so the high and low expected return portfolios from the one-variable regression maximize the spread in the explanatory variable. A new variable can increase the expected return spread only by replacing stocks in the extreme portfolios, and the turnover must reduce the spread in the original variable. This logic, which generalizes to multivariate regressions, says a larger increase in the expected return spread necessarily implies more turnover and more attenuation of the original variables.

Since Fama-MacBeth *t*-statistics measure whether specific variables are related to the cross section of expected returns, one might be tempted to argue that our focus on average return spreads is misplaced. However, a variable can have substantial power to describe variation in average returns when other variables are held constant (the thought experiment captured by the *t*-statistics for average multivariate FM regression slopes) but much less power to produce increments to average return spreads when faced with the joint variation of all variables in a model (the experiment implied by sorts on regression fitted values). To understand how variables combine to produce the expected returns available to investors, it is better to examine spreads in average returns on portfolios formed using regression forecasts with and without the additional explanatory variable.

Two earlier papers use regression fitted values to provide evidence on incremental returns. Fama and French (2006) examine average return spreads for portfolios that split stocks on the median fitted value from cross-section regressions. They comment that adding variables with strong marginal explanatory power in FM regressions (large *t*-statistics for average slopes) produces only modest improvements in average return spreads, but there is a suspicion that this is due to splitting stocks into just two portfolios. Lewellen (2011) finds, however, that adding explanatory variables to his cross-section regressions produces modest improvements in average return spreads for decile portfolios.

Lewellen (2011) tests models that use three (market capitalization, which we call *Size;* the book-to-market ratio *B/M;* and momentum), seven (first three, plus stock issues, accruals, profitability, and asset growth), and 15 (first seven plus eight more) forecasting variables. It is not surprising that the jump from seven to 15 variables produces little gain in average return spreads since many of the additional variables have weak explanatory power in the regressions. The small improvement in the jump from three to seven variables is more surprising because the additional variables show strong explanatory power. Similar comments apply to Fama and French (2006), whose regression models contain two, five, seven, and nine explanatory variables.

Fama and French (2006) and Lewellen (2011) do not explain why variables that have strong explanatory power when added to cross-section regressions produce only small increments to average return spreads in sorts on regression fitted values. That is the task addressed here.

We emphasize that we do not provide an exhaustive examination of how the panoply of average return variables identified in the literature combine to produce the cross section of expected returns. Our more limited goal is to illuminate the two general forces – slope and variable attenuation – that explain why variables that produce large average return spreads in univariate sorts and have strong marginal explanatory power in multivariate cross-section return regressions typically have less impact on the investment opportunity set when faced with competition from other variables. To keep the story simple, we focus on the three explanatory variables of Lewellen's first model (*Size*, *B/M*, and momentum), which are among the premier candidates from the literature.

For more perspective on whether an incremental variable improves an investor's opportunity set, we turn to the GRS statistic of Gibbons, Ross, and Shanken (1989). The GRS statistic is commonly used in time-series tests of asset pricing models. The GRS test asks whether the expected returns on a model's factor portfolios span the expected returns on a broader set of test portfolios.

We use the GRS statistic in an unusual way, to test whether quintile portfolios formed by sorting on the fitted values from two-variable regressions span the opportunity set obtained with quintile portfolios built from regressions that add a third variable. The arguments in GRS (1989) say that this is equivalent to testing whether the highest Sharpe ratio one can construct using the quintile portfolios from both the two- and three-variable regressions is reliably higher than the highest Sharpe ratio one can construct with just the quintiles from the two-variable regression.

We find that adding a third variable produces a large increase in the maximum Sharpe ratio and a strong GRS rejection of the hypothesis that the quintile portfolios formed by sorting on the fitted values from a two-variable regression span the opportunity set that can be obtained with three variables. This evidence is, however, of little use to investors. The portfolios that produce the maximum Sharpe ratios in the GRS statistic typically involve unrealistic leverage, with short positions that commonly exceed one hundred times the portfolio investment. When we examine the maximum Sharpe ratios that can be obtained in the absence of short selling (the opportunity set of a long-only investor), the increments to the ratio obtained by adding the third variable are typically small, often trivial.

We proceed as follows. Section 2 uses a simple model of the return generating process to explain how slope and variable attenuation combine to determine a variable's incremental contribution to a multivariate forecast of returns. Sections 3 and 4 illustrate our conclusions with regressions that use *Size*, *B/M*, and momentum to forecast returns. Section 5 presents evidence on the importance of short sales in the tangency portfolios at the heart of the GRS test and explores the impact of short sale constraints on maximum Sharpe ratios. A summary and conclusions are in Section 6. Appendix A provides additional empirical results and additional formal analysis of the model of Section 2.

2. Incremental return spreads: a simple model

We use fitted values from FM regressions of individual stock returns on characteristics to sort stocks into quintile portfolios based on regression estimates of expected returns. This section presents a simple model that illustrates the roles of slope and variable attenuation on the incremental spread in quintile portfolio expected returns when a variable is added to the forecasting regression.

Suppose returns are generated by the model,

$$R_i = b_1 X_{1i} + b_2 X_{2i} + v_i. \tag{1}$$

In Eq. (1), R_i is the return on stock i, X_{1i} and X_{2i} are characteristics of i, v_i is noise that is uncorrelated with X_{1i} and X_{2i}, and b_1 and b_2 are constants. Dropping the subscript i, we assume without loss of generality that the means of X_1 and X_2 are zero and the variances are one. We also assume X_1 and X_2 are defined so that b_1 and b_2 are positive. We make these assumptions to simplify the presentation; violations do not change our conclusions. Finally, we assume X_1 and X_2 are normally distributed.

2.1. *Expected return spreads from univariate and bivariate regressions*

Suppose X_2 is a recently discovered expected return variable and quintile portfolios were previously formed using sorts on fitted values from the univariate regression of R on X_1,

$$R = b_1^* X_1 + u. \tag{2}$$

The univariate slope b_1^* is determined by the bivariate slopes, b_1 and b_2 in Eq. (1) and the correlation between X_1 and X_2, ρ. Since the means of X_1 and X_2 are zero, the intercept is also zero in the regression of X_2 on X_1, and since the standard deviations of X_1 and X_2 are both one, the slope is ρ,

$$X_2 = \rho X_1 + e. \tag{3}$$

Substituting ρX_1 for X_2 in Eq. (1) shows that X_1's slope in its univariate return regression is $b_1 + \rho b_2$,

$$R = (b_1 + \rho b_2) X_1 + u, \tag{4}$$

and the expected return given X_1 is $(b_1 + \rho b_2) X_1$.

We assume X_1 and X_2 are normally distributed, so the conditional expected returns from the univariate and bivariate regressions, $(b_1 + \rho b_2) X_1$ and $b_1 X_1 + b_2 X_2$, are also normally distributed. Thus, if we sort on conditional expected returns, the spread between expected returns for the top and bottom quintile portfolios is proportional to the standard deviation of the conditional expected returns. This implies that a variable's incremental contribution to the spread is determined by its impact on the standard deviation of the regression fitted values.

It is easy to describe how the variance of the conditional expected return (which is a monotone function of the standard deviation) increases when we add X_2 to X_1's univariate return regression. By construction, e, the error term in regression (3) of X_2 on X_1, is uncorrelated with X_1. Thus, we can write the bivariate return regression as

$$\begin{aligned} R &= b_1 X_1 + b_2 X_2 + v \\ &= (b_1 + \rho b_2) X_1 + b_2 e + v, \end{aligned} \tag{5}$$

and the variance of the bivariate fitted values is

$$(b_1 + \rho b_2)^2 \mathrm{Var}(X_1) + b_2^2 \mathrm{Var}(e) = (b_1 + \rho b_2)^2 + b_2^2 (1 - \rho^2). \tag{6}$$

The first term in Eq. (6) is the variance of the conditional expected return from the univariate regression (4). Adding X_2 to Eq. (4) thus adds $b_2^2(1-\rho^2)$ to the variance of the conditional expected return. By this metric, X_2 contributes a lot if its correlation with X_1 is close to zero and its bivariate slope b_2 is large or, equivalently, if little of its information is already captured by X_1 in the univariate regression and it has lots of power in the bivariate regression. Either a strong correlation (positive or negative) or a small bivariate slope prevents an incremental variable from adding much to the variance of the model's forecasts.

Eq. (6) describes the effect of an incremental variable on the variance of the regression fitted values. The univariate and bivariate expected return spreads, however, are proportional to the standard deviations of the fitted values. Since we assume the standard deviations of X_1 and X_2 are both one, the standard deviation of the univariate fitted values, $(b_1 + \rho b_2) X_1$, is the absolute value of the univariate slope, $|b_1 + \rho b_2|$. The standard deviation of the bivariate fitted values, $b_1 X_1 + b_2 X_2$, is $(b_1^2 + \rho b_1 b_2 \ + b_2^2)^{1/2}$. The ratio of the univariate and bivariate expected return spreads, R_1^* and R', is the ratio of these standard deviations,

$$\frac{R_1^*}{R'} = \frac{|b_1 + \rho b_2|}{(b_1^2 + 2\rho b_1 b_2 + b_2^2)^{1/2}}. \tag{7}$$

We can simplify Eq. (7) by defining r as the ratio of the bivariate slopes, $r \equiv b_2 / b_1$,

$$\frac{R_1^*}{R'} = \frac{|1 + \rho r|}{\left(1 + \rho r + r^2\right)^{1/2}}. \tag{8}$$

Thus, if the standard deviations of X_1 and X_2 are both one, the ratio of the univariate and bivariate expected return spreads is determined by ρ, the correlation between X_1 and X_2, and r, the ratio of their bivariate slopes.

Modifying Eq. (8) is simple for variables that do not have standard deviations of one. We could solve the problem by dividing each variable by its standard deviation, but there is no need. Dividing an independent variable by a constant rescales its regression slope and standard error, but it does not change its t-statistic. Thus, the ratio of the bivariate t-statistics for unadjusted variables equals the ratio of the bivariate slopes X_1 and X_2 would have if their standard deviations were one. This means Eq. (8) continues to hold if we redefine r as the ratio of t-statistics, $r \equiv t(b_2)/t(b_1)$.

2.2. *Variable attenuation and amplification*

An incremental variable increases the expected return spread by changing the stocks in the extreme quintiles. In a one variable model, the high portfolio is the stocks with the highest return forecasts, $(b_1 + b_2 \rho) X_1$. Thus, the high portfolio maximizes (or minimizes if the univariate slope is negative) the average value of X_1. The bivariate sort increases the high portfolio's expected return by replacing extreme X_1 stocks that have low X_2 with less extreme X_1 stocks that have high X_2. The same happens in the low expected return portfolio, but in reverse. Turnover in the extreme quintiles increases the expected return spread by increasing the expected spread in X_2, but the turnover also shrinks the expected spread in X_1. If X_2 has incremental forecast power, the net effect on the expected return spread is positive, but the shrinkage of X_1 in the extreme quintiles usually dampens X_2's incremental contribution.

Define X_1^* as the expected difference between the average values of X_1 for the stocks in the high and low expected return quintile portfolios formed using X_1 alone. Similarly, define X_1' and X_2' as the differences between the expected values of X_1 and X_2 for the stocks in the high and low portfolios formed using both X_1 and X_2. [Asterisks (*) and primes (′) denote univariate and bivariate expected spreads.] Finally, define I_2 as X_2's incremental contribution to the bivariate expected return spread, $R' - R_1^*$,

$$I_2 \equiv R' - R_1^* = b_1 X_1' + b_2 X_2' - (b_1 + \rho b_2) X_1^*. \tag{9}$$

We can rewrite Eq. (9) to emphasize the impact of the incremental variable on the expected spreads of X_1 and X_2 in the extreme portfolios,

$$\begin{aligned} I_2 &= b_2(X_2' - \rho X_1^*) - b_1(X_1^* - X_1') \\ &= b_2 \text{Amplification of } X_2 - b_1 \text{Attenuation of } X_1. \end{aligned} \tag{10}$$

The amplification of X_2 is the difference between the expected spread of X_2 in the bivariate extreme portfolios, X_2', and its expected spread in the univariate portfolios, ρX_1^*. The attenuation of X_1 is the difference between the expected spread of X_1 in the univariate and bivariate extreme portfolios, $X_1^* - X_1'$. The amplification of X_2 increases the expected return spread, but the turnover that amplifies X_2 shrinks X_1's expected spread, which usually reduces X_2's incremental contribution.

If the addition of X_2 produces a big increase in the expected return spread, it must also have a big effect on the stocks in the extreme quintiles and on the expected spread in X_1. Appendix A makes this statement precise. Specifically, if returns conform to the model in Eq. (1), the gain in the expected return spread when X_2 is added to X_1's univariate regression equals the gain in the expected X_1 spread when X_2 is removed from the bivariate regression,

$$\frac{R'}{R_1^*} = \frac{X_1^*}{X_1'}. \tag{11}$$

2.3. Another perspective: the return gap

Eq. (10) uses the incremental variable's impact on the composition of the extreme portfolios to explain its impact on the difference between the extreme portfolios' expected returns. For more perspective on the impact of the incremental variable, we consider the difference between the sum of the quintile portfolio expected return spreads produced by the regressions that use X_1 and X_2 alone and the expected return spread for the regression that uses both. This difference, which we call the return gap, can be expressed as regression slopes times expected spreads in the variables,

$$\begin{aligned} \text{ReturnGap} &\equiv R_1^* + R_2^* - R' \\ &= (b_1 + b_2\rho) X_1^* + (b_2 + b_1\rho) X_2^* - (b_1 X_1' + b_2 X_2'). \end{aligned} \tag{12}$$

Regrouping terms in Eq. (12) allows us to isolate the effects of slope and variable attenuation,

$$\begin{aligned} \text{Return Gap} &= \left[b_2\rho X_1^* + b_1\rho X_2^*\right] + \left[b_1(X_1^* - X_1') + b_2(X_2^* - X_2')\right] \\ &= \text{Effect of Slope Attenuation} \\ &\quad + \text{Effect of Variable Attenuation}. \end{aligned} \tag{13}$$

If X_1 and X_2 are uncorrelated ($\rho = 0.0$), the univariate and bivariate slopes are equal; there is no slope attenuation. But if X_1 and X_2 are correlated, each variable's univariate slope, $b_1 + b_2\rho$ for X_1 and $b_2 + b_1\rho$ for X_2, is its bivariate slope plus a component that reflects the expected return information in the included variable due to its correlation with the excluded variable. Since we assume the bivariate slopes, b_1 and b_2, are positive, if ρ is also positive each univariate slope shrinks by $b_2\rho$ or $b_1\rho$ when the other variable is added to the model. Eq. (13) says that such slope attenuation lowers the bivariate expected return spread by $b_2\rho X_1^* + b_1\rho X_2^*$ and increases the return gap between the sum of the univariate expected return spreads and the bivariate expected spread by the same amount.

If the univariate and bivariate regression slopes are positive but ρ is negative, slope attenuation becomes slope amplification. For example, B/M is negatively correlated with measures of prior return that capture momentum, and we shall see that each variable's positive slope is larger when the other variable is added to the regression. When both univariate slopes are positive, X_1^* and X_2^* are also positive and the slope enhancement due to negative correlation reduces the return gap.

Variable attenuation also contributes to the return gap. The attenuation of X_1, for example, is the difference between its expected spreads in the extreme univariate and bivariate portfolios, $X_1^* - X_1'$. If X_1 and X_2's univariate and bivariate slopes are positive, X_1^* and X_2^* are also positive. Variable attenuation – less extreme spreads for X_1 and X_2 in the bivariate regression – then reduces the bivariate expected return spread and increases the return gap.

Slope and variable attenuation are negatively correlated. When the bivariate slopes b_1 and b_2 are positive and the correlation between X_1 and X_2 is close to one, the univariate slopes, $b_1^* = b_1 + \rho b_2$ and $b_2^* = b_2 + \rho b_1$, exceed the bivariate slopes: each variable attenuates the other's slope in the bivariate regression. But with strong positive correlation between X_1 and X_2 there is little variable attenuation: X_1^* is close to X_1' and X_2^* is close to X_2' because a sort on either is effectively a sort on the other. For lower values of ρ, there is less slope attenuation in the bivariate regression, and when ρ is negative, slope attenuation becomes slope enhancement. But lower correlation between X_1 and X_2 means more sacrifice of the spread in X_1 to increase the spread in X_2, and vice versa: there is more variable attenuation.

Table 1
Average coefficients and t-statistics from monthly regressions of returns on *Size*, *B/M*, and *Prior 2-12*.

We estimate cross-section regressions of individual stock returns for each month t from January 1927 through December 2013 on combinations of *Size*, *B/M*, and *Prior 2-12*. *Size* is minus the natural log of market cap at the end of month $t-1$. The book-to-market ratio, *B/M*, is the log of book equity (B) minus the log of market cap (M) at the end of month $t-1$. *Prior 2-12* is the cumulative return (the sum of the eleven simple monthly returns) from $t-2$ to $t-12$. We estimate regressions for all stocks (All), microcap stocks (Micro, which are NYSE, AMEX, and NASDAQ stocks below the 20th percentile of NYSE market cap at the end of the month $t-1$), small stocks (Small, between the 20th and 50th NYSE percentile), and big stocks (Big, above the 50th NYSE percentile). The table shows averages of the time series of monthly coefficients, and t-statistics for the averages. The results for January 1927 to July 1963 use only NYSE stocks. AMEX stocks first appear in the tests in August 1963, and NASDAQ stocks are added in 1973.

	1927:01-2013:12				1927:01-1963:07				1963:08-2013:12			
Independent variables	All	Micro	Small	Big	All	Micro	Small	Big	All	Micro	Small	Big
Average coefficient												
Size	0.20	0.88	0.03	0.08	0.25	1.35	0.00	0.08	0.17	0.54	0.06	0.09
t-Statistic												
Size	4.74	6.97	0.43	3.00	3.02	4.79	0.01	1.79	3.94	7.45	0.81	2.44
Average coefficient												
B/M	0.44	0.58	0.19	0.13	0.44	0.61	0.24	0.16	0.44	0.56	0.16	0.10
t-Statistic												
B/M	6.02	7.76	2.76	1.89	3.06	4.09	1.99	1.36	6.19	7.90	1.92	1.32
Average coefficient												
Prior 2-12	0.82	0.52	1.00	1.25	1.03	0.55	1.02	1.45	0.66	0.49	0.98	1.10
t-Statistic												
Prior 2-12	5.13	3.08	6.42	6.63	3.35	1.58	3.70	4.50	4.14	3.40	5.48	4.89
Average coefficient												
Size	0.12	0.71	−0.03	0.06	0.14	1.11	−0.13	0.05	0.11	0.42	0.03	0.07
B/M	0.34	0.39	0.20	0.11	0.31	0.35	0.25	0.14	0.36	0.42	0.16	0.09
t-Statistic												
Size	3.52	6.31	−0.50	2.28	2.34	4.48	−0.96	1.25	2.65	5.88	0.48	1.92
B/M	5.42	5.83	2.79	1.63	2.71	2.71	2.04	1.18	5.27	6.14	1.90	1.12
Average coefficient												
Size	0.21	0.90	0.05	0.08	0.26	1.34	0.03	0.07	0.18	0.58	0.06	0.08
Prior 2-12	0.81	0.65	0.99	1.22	0.99	0.60	1.00	1.44	0.68	0.68	0.98	1.07
t-Statistic												
Size	5.36	7.57	0.66	2.79	3.34	5.05	0.23	1.54	4.57	8.42	0.89	2.39
Prior 2-12	5.66	4.12	6.39	6.57	3.60	1.86	3.68	4.51	4.67	4.91	5.45	4.78
Average coefficient												
B/M	0.55	0.68	0.35	0.24	0.51	0.65	0.34	0.23	0.59	0.70	0.36	0.25
Prior 2-12	1.13	0.82	1.22	1.40	1.25	0.72	1.19	1.56	1.04	0.89	1.24	1.29
t-Statistic												
B/M	8.60	9.91	5.70	4.28	3.90	4.63	3.15	2.32	9.93	11.65	5.02	3.84
Prior 2-12	8.37	5.45	8.62	8.01	4.97	2.35	4.77	5.29	7.21	6.76	7.56	6.03
Average coefficient												
Size	0.13	0.73	−0.03	0.04	0.14	1.09	−0.10	0.03	0.12	0.46	0.03	0.05
B/M	0.45	0.49	0.35	0.23	0.38	0.39	0.34	0.22	0.49	0.56	0.35	0.23
Prior 2-12	1.04	0.85	1.21	1.37	1.15	0.71	1.18	1.54	0.96	0.95	1.24	1.25
t-Statistic												
Size	3.83	6.69	−0.38	1.54	2.42	4.59	−0.73	0.71	3.01	6.53	0.39	1.39
B/M	8.27	7.87	5.73	4.08	3.77	3.21	3.21	2.26	8.64	9.35	4.98	3.56
Prior 2-12	8.06	5.79	8.57	7.92	4.65	2.36	4.73	5.28	7.26	7.37	7.53	5.92

The return gap has a second interpretation. The difference between the sum of the univariate expected spreads and the bivariate expected spread is also the difference between each variable's univariate expected return spread and its incremental contribution to the bivariate spread,

$$\begin{aligned}\text{Return Gap } &= R_1^* + R_2^* - R' \\ &= R_1^* - (R' - R_2^*) = R_1^* - I_1 \\ &= R_2^* - (R' - R_1^*) = R_2^* - I_2. \qquad (14)\end{aligned}$$

Notice that X_1 and X_2 have the same incremental gap, $R_1^* - I_1 = R_2^* - I_2$. The two variables can produce very different univariate expected return spreads and incremental contributions to the bivariate spread, but the difference between the univariate spread and the incremental contribution is the same for both. And it is equal to the gap between the sum of the univariate expected return spreads and the bivariate expected return spread.

Finally, the analysis above examines the effect of adding a variable to a univariate regression. Our main results continue to apply when the initial regression includes more than one variable. Some results carry over directly. For example, Appendix A shows that the addition of a variable with power attenuates all informative variables already in the regression, regardless of number. Other results persist if we replace X_1 in the analysis above with the fitted value from a multivariate regression. For example, an incremental variable has little impact on the expected return spread from a multivariate model unless it has substantial power in the full regression and its correlation with the fitted values from the original regression is far from 1.0 and -1.0.

We turn now to empirical examples that illustrate these and our other propositions.

3. The regressions

We estimate cross-section regressions of excess returns (returns minus the one-month T-bill rate) on individual stocks for each month t from January 1927 to December 2013 on combinations of *Size*, *B/M*, and a momentum explanatory return. The analysis above assumes the bivariate regression slopes are positive. Thus, we define *Size* as minus the natural log of market cap at the end of month $t-1$, so its regression slopes are generally positive. Although none of our conclusions depend on this definition, it simplifies the links between our analytical and empirical results.

As in our previous work, to be sure book equity (B) in B/M is known in month t, we update B once a year at the end of June, using book equity for the fiscal year ending in the preceding calendar year. The book-to-market ratio is the log of B minus the log of market cap, M, at the end of month $t-1$, with M adjusted for changes in shares outstanding between the measurement of B and the end of $t-1$. Returns and market cap are from the Center for Research in Security Prices (CRSP). Book equity is typically from Compustat and is the value of stockholders' equity, plus balance sheet deferred taxes and investment tax credits (if available), minus the redemption, liquidation, or par value of preferred stock (in that order). As in Davis, Fama, and French (2002), we fill in book equity for NYSE firms without Compustat data.

The explanatory return for momentum, *Prior 2–12*, is the cumulative excess return (the sum of the eleven monthly excess returns) from $t-2$ to $t-12$. Skipping month $t-1$ is common in the momentum literature. It is motivated by evidence that the negative autocorrelation of monthly individual stock returns at the first monthly lag, studied in detail by Jegadeesh (1990), is at least in part a market microstructure effect (for example, Jegadeesh and Titman, 1995). *Prior 2–12* requires a year of past returns for each regression month t. To have the same sample of stocks for all regressions, we exclude stocks without *Size*, *B/M*, or *Prior 2–12* for month t from all the regressions for t. Stocks with negative B are also excluded.

Table 1 shows average slopes from monthly FM cross-section regressions and t-statistics for the average slopes. (Regression intercepts are not important in our analysis, and we do not report them.) We show results for January 1927 to December 2013 and two subperiods that split in August 1963. The first subperiod covers only NYSE stocks. AMEX stocks start on CRSP in August 1962, but because the momentum variable requires 12 months of past returns, they first appear in the tests in August 1963. NASDAQ stocks are added in 1973. We show results for All stocks, and for Microcaps (NYSE, AMEX, and NASDAQ stocks below the 20th percentile of NYSE market cap at the beginning of regression month t), Small stocks (between the 20th and 50th percentiles), and Big stocks (above the 50th percentile).

Formal tests in Appendix Table A1 suggest that the true regression slopes for *Size* in the Microcap regressions are different for 1927–1963 and 1963–2013. In contrast, there is little evidence that the *Size* slopes for Small and Big stocks and the *B/M* and *Prior 2–12* slopes for all three size groups differ between the two periods. Thus, to simplify the discussion we emphasize the full-period slopes. The tests in Appendix Table A2 say that the true regression slopes differ across the three size groups. The average *Size* and *B/M* slopes are more extreme for Microcaps than for Small and Big stocks, and the average *Prior 2–12* slope is most extreme for Big stocks and least extreme for Microcaps.

Much of the regression evidence in Table 1 is old hat. For example, used alone or in combination with other variables, *Size* and *B/M* have rather strong average slopes in the regressions for 1927–2013 that use All stocks. The positive slopes for both variables say that, as usual, smaller stocks and value (high *B/M*) stocks tend to have higher average returns than bigger stocks and growth stocks. The separate regressions for size groups show that much of the action in *B/M* comes from smaller stocks, especially Microcaps. With notable exceptions discussed below, the average *B/M* slope for Big stocks is less than two standard errors from zero even for the full 1927–2013 period. The smaller value premium for Big stocks is well known (Fama and French, 1993, 2012; Loughran, 1997). The average *Size* slopes are also stronger for Microcaps. The average *Size* slopes for Small and Big stocks are typically about one tenth the average Microcap slopes, and the average *Size* slopes for Small stocks are less than one standard error from zero. Note, however, that partitioning the sample into size groups shrinks the range of *Size* in a regression and so reduces power to detect a size effect.

Motivated by Jegadeesh and Titman (1993), many papers document return momentum. Ours is no exception. Used alone or with other variables, the average slopes for *Prior 2–12* in Table 1 are more than three standard errors above zero for almost all size groups and all periods examined. Microcaps in the first subperiod, with two of four t-statistics below 2.0, produce the only exceptions. The average *Prior 2–12* slopes are, however, tiny. For example, when *Prior 2–12* is the only explanatory variable, the largest average slope (for the Big stocks of 1927–1963) is 1.45 ($t=4.50$). Since the slopes are multiplied by one hundred, this estimate says that on average only 1.45% of the *Prior 2–12* return shows up as a momentum return in month t. We shall see, however, that tiny average slopes imply large spreads in average returns in the extremes of sorts that use them to predict returns.

There is an interesting result in the regressions that combine *B/M* and *Prior 2–12*. The average slope for *B/M* or *Prior 2–12* increases when the other is added to a regression. The 1927–2013 average slope for *B/M*, for example, is 0.13 ($t=1.89$) in the univariate regression for Big stocks and 0.24 ($t=4.28$) in the regression that also includes *Prior 2–12*. What explains this result? Beginning with Fama and French (1992), *B/M* is typically estimated at the end of June each year as the ratio of book equity for the fiscal year

Table 2
Average spreads between monthly returns on high and low expected return quintiles formed on regression fitted values.

For each regression model, stocks are allocated to quintiles using regression fitted values. The fitted values for month t combine the explanatory variables for t with the average slopes for the period and size group under consideration: All, Micro, Small, or Big. The table shows the average spreads between the returns on the high and low quintiles, and t-statstics for the average spreads. Variables are defined in Table 1.

	1927:01-2013:12				1927:01-1963:07				1963:08-2013:12			
Independent variables	All	Micro	Small	Big	All	Micro	Small	Big	All	Micro	Small	Big
Average spread												
Size	1.03	1.59	0.02	0.27	1.21	1.95	−0.04	0.30	0.90	1.34	0.07	0.25
B/M	1.22	1.69	0.51	0.32	1.29	1.93	0.69	0.34	1.16	1.51	0.38	0.31
Prior 2-12	0.81	0.60	0.94	0.74	0.75	0.33	0.90	0.80	0.84	0.79	0.96	0.69
Size, B/M	1.30	1.96	0.51	0.42	1.39	2.43	0.65	0.41	1.25	1.71	0.39	0.38
Size, Prior 2-12	1.20	1.87	0.93	0.77	1.29	2.19	0.90	0.82	1.13	1.70	0.95	0.72
B/M, Prior 2-12	1.54	1.89	1.08	0.87	1.50	1.82	1.03	0.88	1.60	1.93	1.12	0.86
Size, B/M, Prior 2-12	1.65	2.07	1.08	0.86	1.59	2.37	1.06	0.88	1.70	2.18	1.13	0.87
t-Statistic for average spread												
Size	4.68	7.10	0.30	3.12	2.95	4.30	−0.25	1.95	3.81	6.52	0.96	2.50
B/M	5.68	7.59	2.84	1.95	3.01	4.24	2.19	1.19	5.86	7.72	1.81	1.58
Prior 2-12	4.46	2.67	5.25	4.22	2.44	0.74	3.15	2.82	3.88	3.64	4.24	3.15
Size, B/M	5.57	7.61	2.79	2.83	2.99	4.70	2.13	1.49	5.74	8.25	1.86	2.47
Size, Prior 2-12	6.97	9.61	5.27	4.50	4.28	5.36	3.16	2.93	5.63	9.14	4.20	3.37
B/M, Prior 2-12	10.11	10.48	7.38	5.61	5.22	4.68	4.29	3.41	11.87	13.91	6.06	4.48
Size, B/M, Prior 2-12	9.39	9.84	7.37	5.52	4.68	5.19	4.44	3.41	10.97	13.72	6.09	4.59

ending in the previous calendar year divided by market cap at the end of December of the preceding year. This value of *B/M* is then used in the regressions for the 12 months beginning in July. We use the same convention for *B* here, but we update *M* each regression month t using market cap at the end of $t-1$. Updating *M* has two offsetting effects in the regressions that use only *B/M* to explain returns. First, *B/M* is a more up-to-date measure of value or growth, which in itself should increase the *B/M* slope. This is offset by momentum, which predicts, for example, lower returns after recent stock price declines that raise *B/M* and move stocks toward value. Likewise, when *Prior 2–12* is used alone, positive momentum is partly offset by the fact that price increases tend to move stocks toward the growth category and lower average returns. Including both updated *B/M* and *Prior 2–12* as explanatory variables alleviates this conflict, allowing the two variables to focus better on the marginal effects of value and momentum. (Asness and Frazzini, 2013, make a closely related point.)

Stated more succinctly, the monthly updated versions of *B/M* and *Prior 2–12* are negatively correlated (about −0.27), and the average slopes for the two variables are positive in their univariate and bivariate return regressions. The model of Section 2 then tells us that the average slopes in the bivariate regression are enhanced relative to the slopes in the univariate regressions.

Finally, in the Table 1 results for 1927–2013, the average slopes in the three-variable regressions for All stocks and Microcaps are more than 3.8 standard errors from zero. In the three-variable regressions for Small and Big stocks, the average *Size* slopes are less than 1.6 standard errors from zero, but the average slopes for *B/M* and *Prior 2-12* are more than 4.0 standard errors from zero. Thus, *Size*, *B/M*, and *Prior 2-12* have lots of marginal information about average returns. We see next that this marginal information often does not translate into large incremental contributions to average returns.

4. Quintile portfolio returns

To measure the cross section of average returns implied by a regression, we construct equal-weight (EW) quintile portfolios monthly by sorting stocks on projected returns from the regression. We examine results for All stocks and for Microcap, Small, and Big stocks. The projections for month t combine explanatory variables for t with average coefficients for the period and size group under consideration. The average coefficients in the projections for month t include the coefficients for t, but we can report that excluding month t has little effect on the results.

Why do we use average slopes to form portfolios for month t rather than the slopes for t? Average slopes provide estimates of the *Size*, *B/M*, and momentum patterns in expected returns we seek to examine. In contrast, the slopes for month t respond to unexpected returns that overwhelm expected returns. For example, when growth stocks have higher returns than value stocks in month t, the regression for t captures this outcome with a negative *B/M* slope, even though average returns are positively related to *B/M*. If we form portfolios using month-by-month slopes, we observe huge spreads in fitted and actual returns that are mostly due to unexpected returns, not to patterns in expected returns.

In using full-period or subperiod average slopes to form portfolios, we are not examining executable investment strategies. This is in line with our goal: to explain why variables that have strong explanatory power when added to cross-section regressions often produce small increments to average return spreads. We have also examined average returns for quintile portfolios that forecast month

Table 3
Average monthly return spreads, incremental contributions, and variable attenuation, 1927–2013.

In Panel A we use fitted values computed with the average slopes in univariate and bivariate regressions to allocate stocks to quintiles each month. R_1^* and R_2^* are the average spreads between the monthly returns on the high and low quintiles formed on the original and incremental variables and R' is the average spread for the regression that includes both. R_1^*/R' and R_2^*/R' are ratios of the average return spreads. I_2/R' is X_2's average incremental contribution as a fraction of the average bivariate spread, $(R'-R_1^*)/R'$. We use $\rho(X_1, X_2)$, the average correlation of the explanatory variables, and $r=t_2/t_1$, the ratio of the t-statistics for their bivariate slopes, in Eq. (14) to compute Predicted I_2/R', and $\rho=0$ is the value of I_2/R' we would predict if the correlation were zero. X_1'/X_1^* is the ratio of the average X_1 spreads from the bivariate and univariate regression sorts. In Panel B the first explanatory variable is the fitted value (FV) from the indicated first-pass bivariate regression. Variable definitions are in Table 1.

			I_2/R'		Predicted I_2/R'				
	R_1^*/R'	R_2^*/R'	Average	t-Statistic	ρ	$\rho=0$	$r=t_2/t_1$	ρ	X_1'/X_1^*
Panel A: Adding a variable to a univariate regression									
Adding *B/M* to *Size*									
All	0.79	0.94	0.21	3.00	0.26	0.46	1.54	0.37	0.77
Micro	0.81	0.86	0.19	3.24	0.18	0.27	0.92	0.30	0.90
Small	0.05	1.01	0.95	2.79	0.90	0.82	−5.60	0.08	0.00
Big	0.65	0.77	0.35	1.24	0.16	0.19	0.71	0.15	0.69
Adding *Size* to *B/M*									
All	0.94	0.79	0.06	1.24	0.10	0.16	0.65	0.37	0.88
Micro	0.86	0.81	0.14	2.14	0.21	0.32	1.08	0.30	0.72
Small	1.01	0.05	−0.01	−0.24	0.02	0.02	−0.18	0.08	1.00
Big	0.77	0.65	0.23	2.00	0.34	0.42	1.40	0.15	0.84
Adding *Prior 2-12* to *Size*									
All	0.86	0.67	0.14	1.12	0.35	0.31	1.05	−0.10	0.77
Micro	0.85	0.32	0.15	2.28	0.13	0.12	0.54	−0.12	0.93
Small	0.03	1.00	0.97	4.54	0.91	0.90	9.62	−0.01	0.03
Big	0.35	0.96	0.65	2.57	0.59	0.61	2.35	0.03	0.31
Adding *Size* to *Prior 2-12*									
All	0.67	0.86	0.33	1.92	0.31	0.27	0.95	−0.10	0.56
Micro	0.32	0.85	0.68	4.23	0.61	0.52	1.84	−0.12	0.28
Small	1.00	0.03	0.00	−0.14	0.01	0.01	0.10	−0.01	1.00
Big	0.96	0.35	0.04	1.15	0.08	0.08	0.43	0.03	0.97
Adding *Prior 2-12* to *B/M*									
All	0.79	0.52	0.21	1.99	0.39	0.28	0.97	−0.28	0.67
Micro	0.89	0.32	0.11	1.46	0.15	0.12	0.55	−0.26	0.82
Small	0.47	0.87	0.53	2.28	0.65	0.45	1.51	−0.30	0.34
Big	0.37	0.85	0.63	2.12	0.74	0.53	1.87	−0.27	0.23
Adding *B/M* to *Prior 2-12*									
All	0.52	0.79	0.48	3.44	0.41	0.30	1.03	−0.28	0.53
Micro	0.32	0.89	0.68	4.41	0.71	0.52	1.82	−0.26	0.35
Small	0.87	0.47	0.13	1.48	0.21	0.17	0.66	−0.30	0.80
Big	0.85	0.37	0.15	2.04	0.14	0.12	0.53	−0.27	0.88
Panel B: Adding a variable to the fitted value from a bivariate regression									
Adding *Prior 2-12* to FV(*Size*, *B/M*)									
All	0.79	0.49	0.21	2.28	0.39	0.29	1.00	−0.25	0.74
Micro	0.96	0.29	0.04	0.75	0.17	0.14	0.60	−0.21	0.90
Small	0.46	0.85	0.54	2.34	0.65	0.45	1.51	−0.30	0.34
Big	0.48	0.85	0.52	1.86	0.61	0.48	1.65	−0.19	0.31
Adding *B/M* to FV*(Size, Prior 2-12)*									
All	0.77	0.78	0.23	2.86	0.18	0.20	0.75	0.11	0.76
Micro	0.86	0.78	0.14	1.98	0.17	0.22	0.79	0.20	0.84
Small	0.85	0.47	0.15	1.66	0.21	0.16	0.66	−0.30	0.80
Big	0.88	0.37	0.12	1.84	0.11	0.09	0.47	−0.23	0.90
Adding *Size* to FV*(B/M, Prior 2-12)*									
All	0.94	0.63	0.06	1.18	0.04	0.05	0.32	0.25	0.89
Micro	0.88	0.75	0.12	1.82	0.14	0.18	0.71	0.22	0.73
Small	0.99	0.02	0.01	2.05	0.00	0.00	0.02	0.04	1.00
Big	0.99	0.31	0.01	0.42	0.03	0.03	0.24	0.10	0.99

t returns using average regression slopes for the preceding 36 months. This approach has two benefits: it allows for variation through time in the underlying true slopes and it is an investible strategy. There is also a potential cost: average slopes for a shorter period have more measurement error if the true slopes do not vary through time. Skipping the details, we can report that using average slopes for 36 months of prior returns produces smaller spreads in quintile portfolio average returns than using full-period average slopes. Apparently, the benefits of increased precision outweigh any costs of time-varying true slopes.

Appendix Table A3 shows average excess returns for EW quintile portfolios formed using full-period or subperiod average slopes. The regressions used to forecast returns do not favor large or small stocks, so EW portfolios capture the implications of the regressions for average returns. EW results for All stocks can be dominated by Microcaps and Small stocks, which are especially plentiful after 1963, so we also provide separate results for Microcap, Small, and Big stocks. Table A3 shows that sorting stocks into quintiles each month using projections of returns based on average regression slopes typically produces monotone increasing average portfolio returns. This result is not hard-wired. It says, for example, that the regressions probably do not have serious functional form problems. Moreover, we can report that R^2 imputed using average slopes to project returns are close to zero, which means the monthly residual returns produced using average slopes have about the same dispersion as raw returns. Table A3 nevertheless shows that the random variation in monthly returns that results in low R^2 does not prevent the average regression slopes from identifying meaningful patterns in average returns.

4.1. *Average return spreads*

For each of the seven regressions, Table 2 shows average differences between the returns on the EW quintile portfolios predicted to have the highest and lowest average returns. Most of the average return spreads are large. In the 1927–2013 results that use All stocks, the spreads are 0.81% per month (about 9.7% per year) or greater, and all are more than 4.4 standard errors from zero. Microcap stocks typically produce the largest spreads. With one exception, the spreads for Microcaps for 1927–2013 are at least 1.59% per month and 7.10 standard errors from zero. The spreads tend to be less extreme for Small stocks and smaller still for Big stocks, but even for Big stocks the full-period average return spreads are at least 0.27% per month (about 3.2% per year) and 1.95 to 5.61 standard errors from zero.

Five of seven Microcap average return spreads fall from 1927–1963 to 1963–2013, but the 1963–2013 spreads are at least 0.79% per month and 3.64 standard errors from zero. For Small and Big stocks, average spreads are similar for the two subperiods. For example, the subperiod averages from the three-variable regression are 1.06% (t=4.44) and 1.13% (t=6.09) for Small stocks and 0.88% (t=3.41) and 0.87% (t=4.59) for Big stocks. Because of the similarity of average spreads (Table 2) and average regression slopes (Table 1) for the two subperiods, henceforth we focus on the full 1927–2013 period.

Judged on t-statistics for average slopes (Table 1), *Size* and *B*/*M* are generally more important than *Prior 2-12* in the Microcap regressions and *Prior 2-12* is more important for Small and Big stocks. Reinforcing these conclusions, the 1927–2013 quintile portfolio average return spreads from the univariate *Size* and *B*/*M* regressions for Microcaps are larger than the spreads for Small and Big stocks, and the average spread from the univariate *Prior 2-12* regression is larger for Small and Big stocks.

We next show that the average return spreads produced by univariate regressions typically become much smaller incremental contributions to average return spreads from multivariate models. We also examine the relation between observed incremental contributions and those predicted by the model of Section 2.

4.2. *Incremental contributions*

The first two columns of Panel A of Table 3 show R_1^*/R' and R_2^*/R', the spreads in 1927–2013 quintile portfolio average returns produced by each of two variables when used alone to forecast returns, expressed as proportions of R', the average return spread produced by a bivariate return regression that includes both variables. The third column shows $I_2/R'=(R'-R_1^*)/R'$, the increment to R_1^* obtained with the bivariate return regression, again measured relative to R'. Panel A shows results for the six bivariate permutations of *Size*, *B*/*M*, and *Prior 2-12* – adding *B*/*M* to *Size*, adding *Size* to *B*/*M*, etc. Panel B is the same except R_1^* is the average return spread produced by a bivariate return regression and R' is the average return spread from a regression that combines the fitted value from the bivariate regression and the third variable. Panel B is thus concerned with incremental average return spreads produced by each of the three explanatory variables when added to a regression that forecasts returns with the other two. For example, we use projected returns computed with the 1927–2013 average coefficients from the *B*/*M-Prior 2-12* return regression to measure the incremental contribution of *Size*.

The sum of the first two columns in Panel A of Table 3, $(R_1^*+R_2^*)/R'$, is almost always much greater than one. This means the sum of the return spreads the two variables produce individually is much greater than the spread they produce together. The exceptions are when *Size* is used to forecast returns for Small stocks. *Size* has no reliable explanatory power in the return regressions for Small stocks (Table 1), the average Small stock return spreads produced by *Size* alone are trivial (Table 2), and the average return spreads produced by combining *Size* with another variable are close to the spreads produced by the other variable alone (Table 3). The first two columns of Panel B of Table 3 are the average return spreads produced by a bivariate regression and by the third variable's univariate regression, divided by the average return spread produced by a regression that includes the bivariate fitted value and the third variable. As in Panel A, the sum is always much greater than one except when we combine the *Size* and *B*/*M-Prior* 2-12 regressions for Small stocks.

Despite the simplifying assumptions of the model of Section 2, Table 3 shows that its predictions of incremental contributions, I_2/R', line up well with actual contributions. We use the value of R_1^*/R'predicted by Eq. (8) to compute the model's prediction of the incremental contribution as a fraction of the bivariate expected return spread, $I_2/R'=1.0-R_1^*/R'$. The inputs (the determinants of slope and variable attenuation) are the average of the monthly correlations between the explanatory variables, ρ, and r=t_2/t_1, the ratio of the t-statistics for the average slopes for the incremental and initial variables in a bivariate regression model. The correlation between the actual and predicted incremental contributions is 0.93 for Big stocks, 0.99 for Small stocks, and 0.98 for Microcaps. Perhaps because the true regression slopes vary across size groups (Appendix Table A2), the correlation for All stocks is only 0.67.

For perspective on the effects of correlation between explanatory variables on incremental spreads, Table 3 reports what the spreads predicted by the model of Section 2 would be in the absence of correlation ($\rho=0.0$). Section 2 says ρ close to 1.0 or -1.0 dooms a variable to a small incremental contribution, but for the variables used here, correlation is never strong enough to overwhelm what would otherwise be a big incremental contribution. The most extreme correlation is 0.37. The others are between -0.30 and 0.30. When we add *Size* to the *B/M* regression for Microcaps, for example, the ratio of bivariate *t*-statistics is 1.08; that is, the two variables have about the same power in the bivariate return regression. Reducing the correlation of *Size* and *B/M* from the estimated 0.30 to zero increases the predicted incremental contribution of *Size* from 21% of the bivariate spread to 32%, both much less than the average return spread produced by *Size* alone, 81%. In contrast, for Microcaps *B/M* is much more powerful than *Prior 2-12*; the ratio of *t*-statistics is 1.82 when we add *B/M* to the *Prior 2-12* regression for Microcaps; and increasing ρ from the observed -0.26 to zero reduces the large predicted contribution of *B/M* from 71% of the bivariate spread to 51%. (Negative correlation produces slope enhancement that almost always increases incremental contributions when both variables in a bivariate regression have positive average slopes.)

The seventh column of Table 3 reports the ratios of bivariate *t*-statistics, *r*. The ratios signal big differences in incremental contributions. In nine of ten regressions in which the absolute ratio of *t*-statistics exceeds 1.50 (so the relative power of the incremental variable is high), both the predicted and actual incremental contributions exceed 50% of the bivariate spread. In the seven regressions in which the absolute ratio is below 0.50, the predicted and actual incremental contributions are 12% or less.

4.3. Variable attenuation

Attenuation of variables already in the model almost always limits an incremental variable's contribution to the expected return spread produced by a multivariate model. If we start with a univariate model, variable attenuation reduces the incremental contribution unless the new variable reverses the sign of the original variable's slope. If we start with a multivariate model, the fitted value from the original model replaces the original variable, but the concept is the same: Variable attenuation reduces the incremental contribution unless the slope for the projections from the original model changes sign when a new variable is introduced.

The last column of Table 3 shows X'_1/X^*_1, the ratio of X'_1 (the average difference between the average values of X_1 in the extreme quintile portfolios in the sorts using a bivariate return regression) and X^*_1 (the average difference when X_1 is used alone), The correlation between X_1 and X_2 also plays a role, but variable attenuation is clearly related to the relative forecast power of the two variables in a bivariate regression, that is, $r=t_2/t_1$. For Microcaps, for example, *B/M* is much more important than *Prior 2-12* and when *Prior 2-12* is added to the regression, the ratio of *t*-statistics is only 0.55 and the average spread of *B/M* is 82% of the average spread when *B/M* is used alone. In contrast, *Prior 2-12* is much more important than *B/M* ($r=1.87$) for Big stocks, and adding *Prior 2-12* causes the spread in *B/M* to drop to only 23% of its value in the univariate regression.

The model of Section 2 makes a specific prediction about variable attenuation in bivariate regressions: The attenuation of X_1 when X_2 is added to a univariate model matches the reduction in the expected return spread when X_2 is removed from the bivariate model, $X'_1/X^*_1=R^*_1/R'$. The first and last columns of Table 3 show that the actual ratios line up rather well with this prediction. When we add *B/M* to *Size*, for example, the ratio of the average return spreads for All stocks is 0.79 and the matching ratio of average *Size* spreads is 0.77. More generally, the correlation between R^*_1/R' and X'_1/X^*_1 across the nine regression pairs and four size groups is 0.97.

4.4. Combined effects of slope and variable attenuation

Table 4 examines the determinants of the return gap for the three pairs of variables and for each variable paired with the fitted value from a bivariate regression that includes the other two. (To simplify the discussion, we often ignore the distinction between the two sets of results.) The results describe the difference, for example, between the sum of the expected return spreads produced by two univariate regressions and the expected return spread from the bivariate regression that uses both,

$$\text{Return Gap} = R^*_1+R^*_2-R'. \tag{15}$$

Eq. (13) says that in the model of Section 2, the return gap is the sum of slope attenuation, $b_2\rho X^*_1+b_1\rho X^*_2$, and variable attenuation, $b_1(X^*_1-X'_1)+b_2(X^*_2-X'_2)$. In Table 4 we use time series averages of the correlations (ρ), bivariate slopes (b_1 and b_2), and variable spreads to estimate the effects of slope and variable attenuation. The return gap and its two components in Table 4 are expressed as proportions of R', the spread in average returns produced by the relevant bivariate regression.

The slopes for *B/M* and *Prior 2-12* are positive in all regressions (Table 4), and because our definition of *Size* gives smaller stocks higher values, the *Size* slope is positive in all but one regression. The correlation between *Size* and *B/M* is also positive (Table 3), so the two variables share information that helps each predict returns. Since the shared information is in the univariate *Size* and *B/M* slopes, each variable's slope falls when the other variable is added. (The exception is the Small stock regression in which *Size* has no reliable explanatory power.) Focusing on All stocks, for example, the average *Size* slope falls from 0.20 to 0.12 when *B/M* is added and the average *B/M* slope falls from 0.44 to 0.34 when *Size* is added. Table 4 shows that the model of Section 2 predicts that attenuation of the two All stock slopes increases the return gap by 48% of the average return spread from the bivariate regression. In contrast, because the correlation between *Size* and *Prior 2-12* is close to zero, slope attenuation has little impact on

Table 4
Average values of the return gap and its determinants, slope and variable attenuation, 1927 – 2013.

We use fitted values (FVs) computed with the average slopes in univariate and bivariate regressions to allocate stocks to quintiles each month. X_1 is the first variable in a pair or the fitted value from the indicated first-pass regression and X_2 is the second variable. $Gap/R' \equiv [(R_1^* + R_2^*) - R']/R'$ is the difference between the sum of the univariate quintile portfolio average return spreads and the bivariate spread, as a fraction of the bivariate spread. The model of Section 2 predicts that the effect of Slope Attenuation on Gap/R' is $(b_2\rho X_1^* + b_1\rho X_2^*)/R'$ and the effect of Variable Attenuation is $[b_1X_1^* + X_1' + b_2(X_2^* - X_2'/R'$. In these expressions, b_1 and b_2 are the average bivariate slopes, X_1^* *and* X_2^* are the average univariate spreads in X_1 and X_2, X_1' *and* X_2' are the average bivariate spreads, and ρ is the average correlation between X_1 and X_2. Univ denotes a variable's average univariate slope. Ave is the average of the monthly values of Gap/R' and *t*-stat is its *t*-statistic computed using the time series standard error. Variable definitions are in Table 1.

	Gap/R'		X_1 slopes		X_2 slopes		Attenuation effects		
	Ave	*t*-Stat	Univ	b_1	Univ	b_2	Slope	Variable	Total
Size and *B/M*									
All	0.73	5.25	0.20	0.12	0.44	0.34	0.48	0.18	0.67
Microcap	0.68	7.30	0.88	0.71	0.58	0.39	0.45	0.23	0.68
Small	0.06	0.33	0.03	−0.03	0.19	0.20	0.13	−0.07	0.06
Big	0.41	2.33	0.08	0.06	0.13	0.11	0.24	0.21	0.45
Size and *Prior 2-12*									
All	0.53	7.95	0.20	0.21	0.82	0.81	−0.05	0.51	0.46
Microcap	0.17	2.23	0.88	0.90	0.52	0.65	−0.12	0.38	0.27
Small	0.03	0.33	0.03	0.05	1.00	0.99	−0.01	0.05	0.05
Big	0.31	3.41	0.08	0.08	1.25	1.22	0.06	0.22	0.28
B/M and *Prior 2-12*									
All	0.31	6.12	0.44	0.55	0.82	1.13	−0.40	0.65	0.26
Microcap	0.21	2.49	0.58	0.68	0.52	0.82	−0.32	0.51	0.19
Small	0.34	3.23	0.19	0.35	1.00	1.22	−0.50	0.67	0.17
Big	0.22	1.67	0.13	0.24	1.25	1.40	−0.41	0.59	0.19
FV(*Size*, *B/M*) and *Prior 2-12*									
All	0.29	5.62	1.00	1.20	0.82	1.08	−0.32	0.60	0.28
Microcap	0.25	3.28	0.93	1.01	0.52	0.80	−0.26	0.47	0.20
Small	0.31	2.98	1.01	1.80	1.00	1.22	−0.49	0.66	0.17
Big	0.33	2.85	1.06	1.53	1.25	1.33	−0.23	0.50	0.27
FV(*Size*, *Prior 2-12*) and *B/M*									
All	0.54	6.69	0.84	0.74	0.44	0.39	0.16	0.31	0.46
Microcap	0.65	8.35	0.86	0.70	0.58	0.46	0.30	0.27	0.57
Small	0.32	3.05	1.01	1.23	0.19	0.35	−0.48	0.65	0.17
Big	0.24	1.79	0.98	1.07	0.13	0.21	−0.29	0.51	0.21
FV(*B/M*, *Prior 2-12*) and *Size*									
All	0.57	6.00	0.91	0.76	0.20	0.14	0.32	0.22	0.54
Microcap	0.63	8.85	0.94	0.72	0.88	0.73	0.34	0.28	0.62
Small	0.01	0.12	0.95	0.94	0.03	0.01	0.03	0.01	0.04
Big	0.30	3.40	0.95	0.93	0.08	0.05	0.14	0.12	0.26

the predicted return gap for the *Size-Prior 2-12* regressions. The strong negative correlation between *B/M* and *Prior 2-12* enhances each variable's slope when the other is added. Amplification of the bivariate slopes reduces the return gap predicted by the model by between 32% and 50% of the bivariate spread.

The story is basically the same when we use fitted values from bivariate regressions to estimate a variable's incremental contribution to the average return spread produced by the three-variable model. A positive correlation between the variable and the fitted value reduces the bivariate slopes and increases the predicted return gap. A negative correlation has the opposite effect.

The predicted effects of variable attenuation on the return gap in Table 4 are more consistent. In all but one case, variable attenuation increases the predicted gap between the sum of the quintile portfolio expected return spreads produced by the univariate regressions that use X_1 and X_2 alone and the expected return spread from the bivariate regression that uses both. (The *Size-B/M* regression for Small stocks, in which *Size* has essentially no power, is the only exception.) Moreover, when slope enhancement reduces the predicted gap, the reduction is always more than offset by the increase caused by variable attenuation.

We argue earlier that slope and variable attenuation are negatively correlated. The correlation between the estimates of slope and variable attenuation in Table 4 is −0.79. The intuition is straightforward. For example, *B/M* and *Prior 2-12* are negatively correlated. As a result, their bivariate average regression slopes are enhanced relative to their univariate slopes, which in itself increases incremental contributions to the average return spread in the bivariate return regression and so reduces the return gap, $(R_1^* + R_2^*) - R'$. But negative correlation also implies substantial variable attenuation: To get higher *B/M* one

Table 5
GRS statistics testing whether two-variable models explain monthly returns on three-variable quintiles, 1927–2013.

We form quintiles each month by sorting stocks on their fitted values from two- and three-variable regressions and then compute monthly returns on the quintiles, as in Table 2. To compute equal-weight (EW) returns, we allocate 20% of the stocks to each quintile and weight stocks equally in a quintile. To compute value-weight (VW) returns, we allocate 20% of the aggregate market cap to each quintile and weight stocks by market cap in a quintile. We use the GRS statistic of Gibbons, Ross, and Shanken (1989) to test whether the returns on the quintiles from a two-variable model explain the expected returns on the quintiles from the three-variable model. Specifically, we regress the excess returns to three-variable quintiles on the excess returns to the two-variable quintiles, and we use the GRS statistic to test whether the intercepts are indistinguishable from 0.0. The probability of observing a GRS statistic greater than the reported value if the true intercepts are 0.0, labeled *p*-value, is in percent. Variable definitions are in Table 1.

	All		Microcap		Small		Big	
Two-variable model	GRS	*p*-Value	GRS	*p*-Value	GRS	*p*-Value	GRS	*p*-Value
EW								
Size and *B/M*	17.43	0.00	12.24	0.00	15.91	0.00	10.01	0.00
Size and *Prior 2-12*	8.53	0.00	7.25	0.00	5.48	0.01	5.96	0.00
B/M and *Prior 2-12*	4.85	0.02	3.49	0.39	0.54	74.82	0.79	55.76
VW								
Size and *B/M*	6.12	0.00	13.30	0.00	14.53	0.00	7.61	0.00
Size and *Prior 2-12*	4.24	0.08	9.55	0.00	4.67	0.07	4.39	0.06
B/M and *Prior 2-12*	4.71	0.03	0.08	99.56	0.98	42.68	3.11	0.86

must sacrifice lots of momentum, and vice versa, which in itself reduces incremental contributions and increases the return gap. In contrast, the positive correlation of *Size* and *B/M* attenuates their positive bivariate return regression slopes, but positive correlation also implies less sacrifice of *B/M* when *Size* is added to the return regression that includes *B/M*, and vice versa.

Finally, the Table 4 estimates of the implications of slope and variable attenuation for average return spreads are from the model of Section 2. The model's estimates of the return gap, the combined effects of slope and variable attenuation in the last column of Table 4, are, however, closely related to the estimates of the actual return gap in the first column of the table. The correlation between the numbers in the two columns is 0.95. Thus, the model captures most of the variation in observed return gaps.

5. GRS tests: a different perspective

The GRS statistic of Gibbons, Ross, and Shanken (1989) is commonly used to test whether an asset pricing model explains the cross section of expected returns. The prediction of most asset pricing models is that the intercepts are zero in regressions of asset excess returns (relative to the risk-free rate) on a model's factor portfolio returns. Gibbons, Ross, and Shanken (1989) show that one can test this prediction by asking whether the largest Sharpe ratio one can construct using only the right-hand-side (RHS) portfolios in the asset pricing regressions (the factors) is reliably lower than the largest Sharpe ratio one can construct with both the left-hand side (LHS) and the RHS portfolios. In other words, does the mean-variance efficient (MVE) tangency portfolio produced by the RHS portfolios coincide with ("span") the MVE tangency portfolio obtained using both the RHS and LHS portfolios.

We use the GRS statistic and its embedded maximum Sharpe ratios to examine how adding a variable known to be related to average returns contributes to the investment opportunities associated with variables also known to be related to average returns. To have a manageable set of results, we test whether portfolios formed by sorting stocks into quintiles on the predicted returns from our two-variable regression models span the MVE tangency portfolio constructed with quintile portfolios formed using the two- and three-variable models. For example, we test whether adding *B/M* improves the tangency portfolio provided by the combination of *Size* and *Prior 2-12* by testing whether the intercepts are indistinguishable from zero when we regress the excess returns for the three-variable quintiles on the excess returns for the quintiles formed using just *Size* and *Prior 2-12*. We emphasize that we are leaning on the spanning interpretation of the GRS test. Specifically, does the MVE tangency portfolio obtained from quintiles formed on *Size* and *Prior 2-12* span the tangency portfolio that also uses quintiles formed on all three variables?

Table 5 shows GRS results for EW portfolios in which (as in Table 2) stocks are allocated evenly across quintiles and weighted equally within quintiles. We also show results for value-weight (VW) quintiles in which each stock is weighted by its market cap and each quintile holds 20% of aggregate market cap. Much lower turnover makes the VW portfolios more appropriate for investors. To save space we show results only for the full 1927–2013 period.

The GRS results for EW portfolios in Table 5 provide little evidence that, after controlling for *B/M* and *Prior 2-12*, *Size* improves the opportunity set for Small and Big stocks. Information about *Size* does improve the maximum Sharpe ratio for All stocks and Microcaps, and *B/M* and *Prior 2-12* have strong incremental expected return information for all four size groups. The *p*-values for these ten GRS statistics are all less than 0.40% and seven round to 0.00%. The GRS tests for the VW quintiles in Table 5 produce similar results, except *Size* now enhances the tangency portfolio produced by the combination of *B/M* and *Prior 2-12* for Big stocks and it does not enhance the tangency portfolio for Microcaps. Apparently, the stronger role for *Size* traces to very tiny stocks for EW Microcap quintiles and to very large stocks for VW Big quintiles.

Table 6

Annualized Sharpe ratios for portfolios of All stocks and all Micro, Small, and Big stocks, and maximum Sharpe ratios that can be constructed with quintiles from a two-variable model or quintiles from two- and three-variable models, 1927–2013.

Each Sharpe ratio is the average monthly excess return for 1927–2013 divided by the monthly standard deviation. We multiply by the square root of 12 to annualize. The Size group ratios in Panel A are for equal-weight (EW) and value-weight (VW) portfolios of All stocks and all Microcap, Small, and Big stocks. The Bivariate Sharpe ratios are the largest that can be constructed by combining EW or VW quintiles from a two-variable model. The Augmented ratios are the largest that can be constructed by combining the two-variable quintiles with the quintiles from the three-variable model. We report the maximum feasible Sharpe ratios with unlimited short selling in Panel B and no short selling in Panel C. Variable definitions are in Table 1.

	All	Microcap	Small	Big
Panel A: Size group				
Equal weight	0.49	0.51	0.43	0.44
Value weight	0.42	0.44	0.44	0.42

	All		Microcap		Small		Big	
	Bivariate	Augmented	Bivariate	Augmented	Bivariate	Augmented	Bivariate	Augmented
Panel B: Unlimited short selling								
EW								
Size and *B/M*	0.82	1.32	0.90	1.26	0.51	1.10	0.46	0.90
Size and *Prior 2-12*	0.84	1.11	1.12	1.31	0.90	1.07	0.72	0.94
B/M and *Prior 2-12*	1.16	1.29	1.26	1.35	1.07	1.09	0.89	0.92
VW								
Size and *B/M*	0.47	0.76	0.59	1.07	0.53	1.07	0.45	0.81
Size and *Prior 2-12*	0.62	0.80	0.64	0.99	0.89	1.04	0.62	0.80
B/M and *Prior 2-12*	0.62	0.82	1.08	1.08	1.04	1.07	0.71	0.83
Panel C: No short selling								
EW								
Size and *B/M*	0.61	0.70	0.72	0.75	0.48	0.58	0.46	0.59
Size and *Prior 2-12*	0.66	0.70	0.77	0.77	0.56	0.58	0.60	0.60
B/M and *Prior 2-12*	0.68	0.70	0.71	0.75	0.58	0.58	0.60	0.60
VW								
Size and *B/M*	0.45	0.54	0.49	0.54	0.49	0.59	0.44	0.55
Size and *Prior 2-12*	0.53	0.54	0.51	0.54	0.57	0.59	0.56	0.56
B/M and *Prior 2-12*	0.51	0.54	0.57	0.57	0.59	0.59	0.54	0.55

Each of the two maximum squared Sharpe ratios at the core of the GRS statistic is for a tangency portfolio that links the risk-free asset and a minimum variance frontier (in mean-standard deviation space). Table 6 shows the maximum annualized Sharpe ratios that can be constructed by combining the quintiles from a two-variable regression model and the maximum from the combination of those five portfolios and the quintiles from the three-variable model. We use the realized monthly returns for 1927–2013 to compute the maximum Sharpe ratios and multiply by the square root of 12 to annualize. As benchmarks for the Sharpe ratios of the GRS test, Table 6 also shows (in Panel A) the Sharpe ratios for the EW and VW portfolios of All sample stocks and all Microcap, Small, and Big stocks. The ratios for the EW versions of these benchmark portfolios are between 0.43 (Small stocks) and 0.51 (Microcap) and the ratios for the VW version are between 0.42 (All and Big) and 0.44 (Microcap and Small).

The maximum Sharpe ratios available with EW quintiles are much larger than the benchmarks. Except for the *Size-B/M* regressions for Small and Big stocks, the maximums for combinations of two-variable quintiles (the bivariate EW ratios in Panel B of Table 6) are roughly 1.6 to 2.5 times the Sharpe ratio of the EW portfolio for the same size group (Panel A). Moreover, augmenting the EW quintiles from the bivariate models with the quintiles from the three-variable model substantially improves the maximum Sharpe ratio, except for Small and Big stocks when the additional variable is *Size*. The addition of *Prior 2-12* to the combination of *Size* and *B/M*, for example, increases the maximum Sharpe ratio from 0.82 to 1.32 for All stocks, from 0.90 to 1.26 for Microcaps, from 0.51 to 1.10 for Small stocks, and from 0.46 to 0.90 for Big stocks. These improvements are responsible for the large GRS statistics and miniscule *p*-values in Table 5. Because *B/M* and *Prior 2-12* have more predictive power than *Size*, the three-variable quintiles add less to the maximum Sharpe ratio produced by the EW *B/M-Prior 2-12* quintiles, especially for Small and Big stocks. All but one of the Sharpe ratios for VW portfolios are smaller than the ratios for EW portfolios. The increments produced by the three-variable model are still large, however, unless the third variable is *Size* and the estimates are for a specific size group.

Unfortunately, the investment implications of the large GRS statistics in Table 5 and the large Sharpe ratios in Panel B of Table 6 are limited, at best. The GRS test assumes investors can sell short with full use of the proceeds. Many of the portfolios that produce the maximum Sharpe ratios require high levels of short selling, so this assumption is not benign. Table 7 shows the leverage of the two tangency portfolios of each GRS test, that is, the total value of the short positions in a tangency portfolio divided by the total value of the portfolio. An investor who puts one dollar in the tangency portfolio of EW quintiles

Table 7
Leverage ratios for tangency portfolios constructed with quintiles from a two-variable model or quintiles from two- and three-variable models, 1927–2013.
A portfolio's leverage ratio is the total value of the short positions in the portfolio divided by the total value of the portfolio. A Bivariate ratio is the leverage for the tangency (maximum Sharpe ratio) portfolio constructed by combining the equal-weight (EW) or value-weight (VW) quintiles from a two-variable model. The Augmented ratio is the leverage for the tangency portfolio constructed by combining the EW or VW quintiles from a two-variable model and a three-variable model. The EW and VW quintiles for the two- and three-variable models are described in Table 6. Variable definitions are in Table 1.

	All		Microcap		Small		Big	
	Bivariate	Augmented	Bivariate	Augmented	Bivariate	Augmented	Bivariate	Augmented
EW								
Size and *B/M*	11.90	1997.60	11.00	316.82	1.23	114.42	1.50	97.72
Size and *Prior 2-12*	9.70	8398.12	15.76	81.32	2.21	201.76	2.34	365.29
B/M and *Prior 2-12*	15.16	2131.06	9.83	311.15	3.29	50.51	4.90	610.12
VW								
Size and *B/M*	1.78	43.67	2.43	195.29	1.74	9.74	1.42	32.45
Size and *Prior 2-12*	4.20	49.98	3.36	109.23	2.00	39.29	1.47	51.59
B/M and *Prior 2-12*	1.70	20.66	6.12	23.07	3.06	134.66	1.71	43.98

from the *Size-B/M* regression for All stocks, for example, would short assets worth $11.90 and use the proceeds to purchase assets worth $12.90.

All the leverage ratios for tangency portfolios constructed with quintiles from the two-variable models exceed 1.4 (140%) and many are greater than 4.0. The tangency portfolios that combine two- and three-variable quintiles are much more levered. The lowest leverage ratio is 9.74, half are over one hundred, and seven are over three hundred. Even modest short selling costs would prevent investors from realizing anything close to the Sharpe ratios that drive the GRS statistics in Table 5.

For perspective on the importance of short sales in the GRS tests, Panel C of Table 6 shows the maximum Sharpe ratios available to investors who cannot sell short. The candidate investments are the same EW and VW expected return quintiles used to compute the GRS statistics in Table 5 and the maximum unconstrained Sharpe ratios in Panel B of Table 6. There are several notable results. First, the maximum ratios typically fall a lot when short sales are eliminated. Many Sharpe ratios shrink by more than 0.30 and eleven of the 48 decline by more than 0.50. Second, the short sale constraint has more impact on the tangency portfolios that combine the two- and three-variable quintiles. This is not surprising since the unconstrained versions of these tangency portfolios use more short selling than the unconstrained two-variable tangency portfolios. (The impact of the short-sale constraint, however, is not simply a function of unconstrained leverage.) Third, without short selling, the Sharpe ratio for the tangency portfolio of expected return quintiles from a *Size-B/M* regression is usually not much higher than the ratio for the matching benchmark portfolio. The maximum ratio formed using EW or VW quintiles, for example, is only 0.05 higher than the matching Sharpe ratio for Small stocks and only 0.02 higher for Big stocks. The *Size-B/M* model produces big increases only for EW quintiles of All stocks (from 0.49 to 0.61) and Microcaps (from 0.51 to 0.72). Skipping the details, we can report that in all but one case, the average return obtained with optimal use of the *Size* and *B/M* patterns in average returns is larger, often substantially larger than that of the benchmark EW or VW portfolio of the size group. The standard deviation, however, is correspondingly higher.

Most important, with unrestricted short selling, adding the quintile portfolios from the three-variable regressions to the quintiles from the two-variable regressions almost always produces a dramatic increase in the maximum Sharpe ratio (Table 6, Panel B). In the absence of short selling, however, the increments to the maximum Sharpe ratios provided by the three-variable quintiles are minor, often nonexistent (Table 6, Panel C). The largest increments, obtained by adding *Prior 2-12* to the combination of *Size* and *B/M*, are 0.13 or less. The potential improvements are smaller for long-only investors who start with the tangency portfolios produced by the *Size-Prior 2-12* quintiles or the *B/M-Prior 2-12* quintiles. No increase in the maximum Sharpe ratio exceeds 0.04 and more than half round to 0.00.

In sum, the importance of short positions in the tangency portfolios at the heart of the GRS statistic means that the test is largely uninformative as a guide for investors who face short-sale constraints and short selling costs. We emphasize, however, that this conclusion is not a criticism of the GRS statistic as a test of asset pricing models. For judging asset pricing models, we want a statistic that ferrets out model problems in an unconstrained way. Our point is that strong model rejections on the GRS test may be inconsequential for investors who cannot sell short freely.

6. Conclusions

Much asset pricing research is a search for variables that improve understanding of the cross section of expected returns. Researchers often claim success if a sort on their candidate produces a large spread in average returns. A better measure, however, is the variable's incremental contribution to the return spread produced by a model that includes the variables already known to predict returns. The incremental impact is almost certain to be substantially smaller than the univariate spread. There are two reasons. First, a candidate with significant forecast power is probably positively correlated with the

Table A1
Comparison of regression slopes across periods, 1927–1963 versus 1963–2013.

This table tests whether the average regression slopes in Table 1 are different for August 1963 to December 2013 and January 1927 to July 1963. The table shows the differences between the average regression slopes for August 1963 to December 2013 and January 1927 to July 1963, and the *t*-statistics for the average differences. The table also shows χ^2 statistics and their *p*-values for tests of the hypothesis that the average slopes in a regression for the later period jointly differ from those for the earlier period. Variable definitions are in Table 1.

	Difference between average slopes				*t*-Statistics			
Explanatory variables	All	Micro	Small	Big	All	Micro	Small	Big
Size	0.08	0.81	−0.05	−0.00	0.90	2.79	−0.33	−0.03
B/M	0.00	0.05	0.09	0.06	0.00	0.30	0.58	0.42
Prior 2-12	0.38	0.05	0.03	0.35	1.08	0.14	0.10	0.88
Size	0.03	0.69	−0.16	−0.02	0.38	2.67	−1.07	−0.41
B/M	−0.04	−0.08	0.09	0.05	−0.33	−0.52	0.61	0.36
Size	0.08	0.75	−0.03	−0.01	0.89	2.76	−0.16	−0.21
Prior 2-12	0.31	−0.09	0.03	0.37	1.00	−0.25	0.08	0.95
B/M	−0.08	−0.05	−0.01	−0.01	−0.59	−0.32	−0.08	−0.11
Prior 2-12	0.21	−0.17	−0.05	0.26	0.72	−0.49	−0.17	0.72
Size	0.02	0.64	−0.12	−0.02	0.28	2.56	−0.83	−0.44
B/M	−0.11	−0.16	−0.01	−0.01	−0.95	−1.21	−0.08	−0.10
Prior 2-12	0.19	−0.24	−0.06	0.29	0.67	−0.73	−0.19	0.80

	χ^2 test				*p*-Value			
Explanatory variables	All	Micro	Small	Big	All	Micro	Small	Big
Size	3.90	125.66	0.64	0.00	4.82	0.00	42.47	95.80
B/M	0.00	0.48	1.09	0.60	99.71	48.91	29.66	43.95
Prior 2-12	5.59	0.13	0.03	2.37	1.81	71.48	85.50	12.40
Size and *B/M*	0.69	92.61	5.64	0.61	70.72	0.00	5.97	73.57
Size and *Prior 2-12*	9.79	120.10	0.18	2.77	0.75	0.00	91.36	25.00
B/M and *Prior 2-12*	3.55	2.67	0.12	1.57	16.98	26.34	94.04	45.65
Size, *B/M*, and *Prior 2-12*	5.31	82.15	3.79	2.22	15.05	0.00	28.52	52.81

forecasted returns from the best previous model. If so, some of the candidate's information about expected returns is already in the model. Second, adding a variable with marginal explanatory power always attenuates the values of other explanatory variables in the extremes of a sort on regression fitted values. This variable attenuation almost always reduces the contribution of the additional variable to the expected return spread. In short, to take advantage of the information in a new variable, some of the information in the original variables must be sacrificed.

The workhorse GRS statistic is a powerful tool for testing asset pricing models, but it is less useful for investment purposes. The problem is that the Sharpe ratios at the heart of the statistic typically involve extreme short positions that are not available to investors. As a result, strong rejections on the GRS test may say little or nothing about whether an incremental variable can be used to improve investment decisions.

Appendix A

A.1. Comparison of regression coefficients across periods and across size groups

Table A1 reports differences (and *t*-statistics for the differences) between the average estimates of the 1926–1963 and 1963–2013 slopes from our seven regressions. It also reports χ^2 tests of the hypothesis that a regression's true slopes are the same in the two periods. (The intercepts are not included in the tests.) The variance-covariance matrix used to calculate a χ^2 statistic is the sum of the first and second period variance-covariance matrices computed using the monthly regression estimates. The *p*-value is the probability that a χ^2 statistic exceeds the observed value under the null hypothesis that the true slopes do not change.

Table A2 compares the regression coefficients for each pair of size groups. The first panel shows average differences between the estimated coefficients for the three combinations of Microcap, Small, and Big stocks for the full period and the two subperiods, as well as *t*-statistics testing whether the true differences are zero. Each T^2 statistic in the Panel B then tests whether the true differences for all the coefficients in a regression are zero. The *p*-value is the probability that a T^2 statistic exceeds the observed value under the null hypothesis that the true differences are zero.

A.2. Average monthly returns for quintiles formed on regression fitted values

Many of our tests use EW returns for quintiles of NYSE, AMEX, and NASDAQ firms formed monthly on fitted values from seven cross-section regressions. Table A3 shows average monthly returns for the quintiles for 1927–2013, 1927–1963, and 1963–2013. High is the quintile with the highest fitted values and Low is the quintile with the lowest fitted values.

Table A2
Comparison of regression slopes for different size groups.

Panel A shows differences between the average regression slopes in Table 1 for two size groups (Microcap versus Small, Microcap versus Big, or Small versus Big) and t-statistics (in parentheses) for the differences. Panel B shows T^2 statistics and p-values for tests of the hypothesis that the average slopes in a regression for one size group jointly differ from those for another size group. Variable definitions are in Table 1.

	Microcap versus Small			Microcap versus Big			Small versus Big		
	1927–2013	1927–1963	1963–2013	1927–2013	1927–1963	1963–2013	1927–2013	1927–1963	1963–2013
Panel A: Differences between average regression slopes for size groups									
Size	0.84	1.35	0.48	0.79	1.26	0.45	−0.05	−0.08	−0.03
	(6.72)	(4.90)	(5.86)	(6.49)	(4.69)	(5.86)	(−0.77)	(−0.62)	(−0.46)
B/M	0.39	0.37	0.40	0.46	0.45	0.46	0.07	0.08	0.06
	(6.14)	(2.96)	(6.53)	(6.69)	(3.38)	(6.84)	(1.75)	(1.16)	(1.34)
Prior 2-12	−0.48	−0.47	−0.49	−0.73	−0.90	−0.61	−0.25	−0.43	−0.12
	(−3.25)	(−1.53)	(−3.88)	(−4.18)	(−2.65)	(−3.49)	(−2.33)	(−2.22)	(−0.99)
Size	0.74	1.23	0.39	0.65	1.06	0.35	−0.10	−0.18	−0.04
	(5.99)	(4.56)	(4.68)	(5.70)	(4.28)	(4.57)	(−1.50)	(−1.44)	(−0.56)
B/M	0.20	0.10	0.27	0.28	0.21	0.34	0.09	0.11	0.07
	(3.21)	(0.81)	(4.74)	(4.27)	(1.59)	(5.27)	(2.28)	(1.56)	(1.69)
Size	0.85	1.31	0.52	0.83	1.27	0.50	−0.03	−0.04	−0.02
	(7.00)	(4.90)	(6.50)	(7.13)	(5.00)	(6.72)	(−0.42)	(−0.28)	(−0.33)
Prior 2-12	−0.34	−0.41	−0.29	−0.58	−0.84	−0.38	−0.23	−0.43	−0.09
	(−2.43)	(−1.40)	(−2.44)	(−3.44)	(−2.58)	(−2.31)	(−2.22)	(−2.26)	(−0.77)
B/M	0.33	0.30	0.34	0.43	0.41	0.45	0.11	0.11	0.11
	(5.51)	(2.61)	(5.96)	(6.89)	(3.35)	(7.27)	(2.75)	(1.51)	(2.50)
Prior 2-12	−0.40	−0.47	−0.36	−0.59	−0.84	−0.41	−0.18	−0.37	−0.05
	(−2.93)	(−1.66)	(−2.94)	(−3.58)	(−2.70)	(−2.36)	(−1.69)	(−1.91)	(−0.40)
Size	0.75	1.19	0.43	0.69	1.07	0.41	0.06	0.12	0.02
	(6.24)	(4.54)	(5.30)	(6.32)	(4.54)	(5.39)	(1.02)	(1.02)	(0.34)
B/M	0.14	0.05	0.20	0.26	0.17	0.32	0.12	0.12	0.12
	(2.40)	(0.43)	(3.72)	(4.27)	(1.47)	(5.20)	(3.11)	(1.70)	(2.83)
Prior 2-12	−0.36	−0.47	−0.29	−0.53	−0.83	−0.31	−0.16	−0.36	−0.02
	(−2.70)	(−1.70)	(−2.44)	(−3.30)	(−2.74)	(−1.84)	(−1.53)	(−1.92)	(−0.15)
Panel B: T^2 statistics (and p-values) comparing average regression slopes for size groups									
Size	45.23	23.97	34.32	42.08	22.01	34.39	0.59	0.39	0.21
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(44.10)	(53.53)	(64.70)
B/M	37.73	8.76	42.66	44.77	11.43	46.84	3.05	1.34	1.80
	(0.00)	(0.32)	(0.00)	(0.00)	0.08	(0.00)	(8.09)	(24.71)	(17.97)
Prior 2-12	10.58	2.35	15.06	17.50	7.04	12.21	5.44	4.94	0.97
	(0.12)	(12.58)	(0.01)	(0.00)	(0.83)	(0.05)	(1.98)	(2.68)	(32.40)
Size, B/M	19.24	9.53	18.19	20.65	8.38	19.49	5.96	3.22	3.19
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.05)	(2.25)	(2.35)
Size, Prior 2-12	18.60	9.96	16.29	20.40	10.64	15.88	4.65	3.09	2.04
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.31)	(2.71)	(10.77)
B/M, Prior 2-12	17.07	7.42	16.42	24.73	8.99	22.73	6.80	3.23	5.21
	(0.00)	0.01	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(2.25)	(0.15)
Size, B/M, 2-12	15.32	7.33	15.56	20.07	8.14	21.27	6.25	3.36	4.07
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(1.01)	(0.29)

A.3. Model extensions

A.3.1. Relaxing assumptions

To simplify the exposition, we assume the means of R, X_1, and X_2 are zero, the standard deviations of X_1 and X_2 are one, and their bivariate slopes are positive. Relaxing these assumptions is easy. In the text we show that, whether the independent variables are standardized or not, we can use the ratio of the bivariate t-statistics to compute r, the ratio of the standardized slopes. Our results are also unaffected if we adjust for nonzero means by subtracting their expected values from R, X_1, and X_2.

The assumption that b_1 and b_2 are positive has almost no impact on our conclusions. It is easy to redefine the independent variables to produce positive bivariate slopes. For example, we measure *Size* as *minus* the natural log of price times shares outstanding to make the links with the analysis here as direct as possible. But such adjustments are unnecessary. In equations in which the sign could matter, each bivariate slope is multiplied or divided by either the expected value of the independent variable or the correlation between the variables. A variable redefinition that reverses the sign of the slope also reverses the sign of these parameters, so it has no net effect on the equations.

Table A3
Average equal-weight monthly returns for quintiles formed on regression fitted values.

For each regression model, 20% of the stocks are allocated to each quintile each month using the regression fitted values. The fitted values for month t combine the explanatory variables for t with the average slopes for the period and size group under consideration: Microcap, Small, or Big. The results for all stocks (All) are constructed using the separate average slopes for the three size groups. The average slopes used in the projections for month t exclude the slopes for t. The header in the first column of each block shows the explanatory variables in the regression, e.g., *Size* in the first block and *Size*, *B/M*, and *Prior 2-12* in the last block. Variable definitions are in Table 1.

	1927:01-2013:12				1927:01-1963:07				1963:08-2013:12			
	All	Microcap	Small	Big	All	Microcap	Small	Big	All	Microcap	Small	Big
Size												
High	1.75	2.70	0.99	0.89	2.09	3.57	1.15	1.10	1.50	2.08	0.88	0.73
2	0.94	1.49	1.00	0.85	1.26	2.50	1.26	1.07	0.71	0.75	0.81	0.69
3	0.90	1.17	1.03	0.81	1.11	1.75	1.35	1.00	0.75	0.75	0.80	0.67
4	0.89	1.08	0.94	0.73	1.03	1.51	1.13	0.94	0.78	0.76	0.80	0.58
Low	0.72	1.11	0.97	0.61	0.88	1.62	1.19	0.81	0.61	0.74	0.80	0.48
B/M												
High	1.80	2.41	1.25	1.00	2.12	3.14	1.51	1.19	1.57	1.87	1.07	0.85
2	1.15	1.72	1.10	0.86	1.36	2.42	1.29	1.08	1.01	1.21	0.97	0.69
3	0.92	1.59	0.95	0.71	1.12	2.50	1.24	0.94	0.77	0.93	0.73	0.55
4	0.73	1.12	0.88	0.66	0.95	1.68	1.20	0.85	0.58	0.72	0.64	0.52
Low	0.59	0.71	0.74	0.67	0.82	1.21	0.82	0.86	0.41	0.35	0.69	0.54
Prior 2-12												
High	1.49	1.79	1.39	1.19	1.68	2.24	1.56	1.42	1.36	1.47	1.27	1.02
2	1.20	1.69	1.15	0.88	1.42	2.35	1.35	1.13	1.04	1.21	1.01	0.70
3	0.99	1.53	1.05	0.71	1.26	2.33	1.35	0.93	0.80	0.96	0.83	0.56
4	0.82	1.33	0.87	0.66	1.08	2.12	1.14	0.84	0.63	0.76	0.68	0.53
Low	0.69	1.20	0.46	0.45	0.93	1.91	0.66	0.62	0.51	0.68	0.31	0.33
Size and *B/M*												
High	1.89	2.82	1.26	1.02	2.22	3.82	1.53	1.21	1.66	2.15	1.07	0.87
2	1.05	1.46	1.10	0.85	1.29	2.25	1.28	1.11	0.89	0.94	0.94	0.68
3	0.91	1.30	0.93	0.74	1.07	1.77	1.17	0.95	0.76	0.78	0.77	0.60
4	0.77	1.11	0.88	0.68	0.97	1.74	1.21	0.86	0.62	0.76	0.63	0.51
Low	0.59	0.86	0.75	0.61	0.83	1.39	0.88	0.80	0.41	0.44	0.68	0.49
Size and *Prior 2-12*												
High	1.74	2.73	1.38	1.21	2.04	3.62	1.53	1.44	1.52	2.06	1.27	1.04
2	1.24	1.44	1.17	0.88	1.45	2.14	1.40	1.14	1.08	1.13	1.01	0.71
3	0.95	1.32	1.07	0.73	1.16	2.00	1.36	0.91	0.78	0.79	0.86	0.58
4	0.73	1.20	0.86	0.65	0.97	1.77	1.14	0.82	0.56	0.73	0.65	0.51
Low	0.54	0.86	0.45	0.44	0.75	1.43	0.62	0.62	0.39	0.36	0.31	0.31
B/M and *Prior 2-12*												
High	1.82	2.43	1.45	1.21	2.06	3.02	1.65	1.42	1.65	1.96	1.31	1.05
2	1.32	1.92	1.28	0.91	1.56	2.79	1.58	1.17	1.16	1.31	1.06	0.74
3	1.02	1.48	1.03	0.80	1.24	2.13	1.22	0.99	0.85	1.05	0.89	0.62
4	0.76	1.18	0.80	0.64	0.93	1.80	1.00	0.80	0.63	0.73	0.65	0.55
Low	0.28	0.53	0.37	0.34	0.57	1.21	0.62	0.54	0.04	0.03	0.19	0.19
Size, *B/M*, and *Prior 2-12*												
High	1.94	2.69	1.46	1.21	2.21	3.58	1.67	1.43	1.76	2.21	1.30	1.06
2	1.28	1.75	1.26	0.91	1.48	2.34	1.52	1.17	1.15	1.25	1.08	0.74
3	0.96	1.36	1.03	0.80	1.10	2.08	1.23	0.99	0.82	0.90	0.88	0.65
4	0.71	1.11	0.79	0.62	0.96	1.75	1.04	0.78	0.55	0.67	0.66	0.52
Low	0.30	0.63	0.38	0.35	0.62	1.21	0.60	0.55	0.06	0.04	0.18	0.18

A.3.2. The attenuation of explanatory variables

If we maintain the assumptions that all random variables are normally distributed, the means of R, X_1 and X_2 are zero, and the standard deviations of X_1 and X_2 are one, we can derive an analytical expression for the attenuation of X_1 when X_2 is added to the regression. Standard deviations of one imply the variance of the residual e in Eq. (3), the regression of X_2 on X_1, is $1-\rho^2$. Define $w \equiv e/(1\text{-}\rho^2)^{1/2}$ as the standardized residual, with a variance of one. By construction, w is uncorrelated with X_1, so X_1's slope in the univariate return regression, $b_1+b_2\rho$, does not change if we add w to the regression,

$$R=(b_1+b_2\rho)X_1+b_2\left(1-\rho^2\right)^{1/2}w+u. \tag{16}$$

Let $C_1^*$and -C_1^* be the minimum and maximum returns in the high and low extreme portfolios formed using X_1 alone, and let C' and -C' be the boundary returns for the extreme portfolios formed using X_1 and X_2. If the extreme portfolios are quintiles, for example, the high X_1 portfolio is the 20% of stocks whose univariate expected return exceeds C_1^*,

$$E(R) = (b_1 + b_2\rho)X_1 > C_1^*, \tag{17}$$

and the bivariate high portfolio is the 20% whose bivariate expected return exceeds C',

$$E(R) = (b_1 + b_2\rho)X_1 + b_2\left(1-\rho^2\right)^{1/2} w > C'. \tag{18}$$

Since X_1 and w are uncorrelated normal (0,1) random variables, their bivariate density function is concentric circles around the origin in a plane defined by X_1 and w. Thus, the expected values of X_1 and w in the bivariate high and low expected return portfolios must be on the line that goes through the origin and is perpendicular to the bivariate expected return boundaries $E(R)=C'$ and $E(R)=-C'$,

$$0 = b_2(1-\rho^2)^{1/2} X_1 - (b_1 + b_2\rho)w. \tag{19}$$

This implies the spreads in the expected values of X_1 and w are proportional to their slopes in the return Eq. (16),

$$w' = \frac{b_2\left(1-\rho^2\right)^{1/2}}{b_1 + b_2\rho} X_1'. \tag{20}$$

The larger the slope on w relative to the slope on X_1 in Eq. (16), the larger the expected spread of w relative to the expected spread of X_1 in the extreme bivariate portfolios.

The expected univariate return spread is the expected univariate spread in X_1 times the univariate slope and the expected bivariate return spread is the sum of the expected bivariate spreads in X_1 and w times their slopes. Thus, the ratio of the expected return spreads is

$$\frac{R_1^*}{R'} = \frac{(b_1 + b_2\rho)X_1^*}{(b_1 + b_2\rho)X_1' + b_2(1-\rho^2)^{1/2} w'}. \tag{21}$$

Eq. (20) allows us to rewrite Eq. (21) as

$$\begin{aligned}\frac{R_1^*}{R'} &= \frac{(b_1 + b_2\rho)^2}{(b_1 + b_2\rho)^2 + b_2^2(1-\rho^2)} \frac{X_1^*}{X_1'} \\ &= \frac{(b_1 + b_2\rho)^2}{b_1^2 + 2\rho b_1 b_2 + b_2^2} \frac{X_1^*}{X_1'} \\ &= \left(\frac{R_1^*}{R'}\right)^2 \frac{X_1^*}{X_1'},\end{aligned} \tag{22}$$

where the last line follows from fact that the ratio of expected return spreads is the ratio of the standard deviations of the univariate and bivariate expected returns, in Eq. (7).

Eq. (22) implies the ratio of the univariate and bivariate expected X_1 spreads is the ratio of the bivariate and univariate expected return spreads,

$$\frac{X_1^*}{X_1'} = \frac{R'}{R_1^*}. \tag{23}$$

When X_2 is added to X_1's univariate regression, the spread between the expected values of X_1 in the high and low portfolios shrinks by the same proportion that the spread in expected returns grows. The conclusion that the spread in X_1 shrinks when X_2 is added to its regression generalizes to the multivariate case: Regardless of the number of explanatory variables in a regression, the addition of another informative variable must shrink the expected spread of every informative variable already in the regression.

Consider a sequence of explanatory variables, X_1, X_2, X_3, and so on. As above, we assume the variables are normally distributed, with means of zero and variances of one. Define U_i as the part of X_i that is orthogonal to all prior variables,

$$X_i = \lambda_{i1}X_1 + \lambda_{i2}X_2 + \ldots + \lambda_{i,i-1}X_{i-1} + U_i. \tag{24}$$

Adding a subscript to identify the number of variables in the model, we write the one-, two-, and three-variable return regressions as

$$R = b_{11}X_1 + v_1 \tag{25}$$

$$R = b_{12}X_1 + b_{22}X_2 + v_2 \tag{26}$$

and

$$R = b_{13}X_1 + b_{23}X_2 + b_{33}X_3 + v_3. \tag{27}$$

To understand how adding X_3 to the two-variable regression affects the extreme portfolios, we replace X_2 and X_3 in Eq. (27) by their orthogonalized versions,

$$\begin{aligned}R &= b_{12}X_1 + b_{22}X_2 + b_{33}U_3 + v_3, \\ &= b_{11}X_1 + b_{22}U_2 + b_{33}U_3 + v_3.\end{aligned} \tag{28}$$

The fitted values from Eqs. (27) and (28) are identical so orthogonalization has no effect on the model's allocation of stocks to the extreme portfolios. The two-variable model uses the return predicted by X_1 and X_2, either $b_{12}X_1 + b_{22}X_2$ or its equivalent, $b_{11}X_1 + b_{22}U_2$, to sort stocks into the high and low expected return portfolios. The three-variable model uses the return predicted by X_1, X_2, and X_3, $b_{13}X_1 + b_{23}X_2 + b_{33}X_3$ or, equivalently, $b_{11}X_1 + b_{22}U_2 + b_{33}U_3$. If X_3 improves the model's forecasts (b_{33} is not zero), the stocks in the extreme portfolios from the three-variable model have more extreme values of X_3 and less extreme values of the two-variable forecast than the stocks in the two-variable portfolios; the expected spread in the two-variable forecasts shrinks when we add a third informative variable to the model. And because X_1, U_2, and U_3 are uncorrelated, if X_1 and U_2 predict returns (b_{12} and b_{22} are both nonzero), both variables contribute to the reduction in the expected spread of the two-variable forecasts. Finally, the expected spread in U_2 cannot shrink unless the expected spread in X_2 also shrinks, so if X_1, X_2, and X_3 predict returns, the addition of X_3 to the two-variable regression must shrink the expected spreads in X_1 and X_2.

There is nothing special about two- and three-variable regressions. Our argument applies when an incremental variable is added to a regression with any number of variables. If the new variable has explanatory power, the expected spread for each of the original variables must shrink.

References

Asness, C., Frazzini, A., 2013. The devil in HML's details. Journal of Portfolio Management 39, 49–68.

Davis, J.L., Fama, E.F., French, K.R., 2002. Characteristics, covariances, and average returns: 1929 to 1997. Journal of Finance 55, 389–406.

Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.
Fama, E.F., French, K.R., 2006. Profitability, investment, and average returns. Journal of Financial Economics 82, 491–518.
Fama, E.F., French, K.R., 2012. Size, value, and momentum in international stock returns. Journal of Financial Economics 105, 457–472.
Fama, E.F., MacBeth, J., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81, 607–636.
Gibbons, M.R., Ross, S.A., Shanken, J., 1989. A test of the efficiency of a given portfolio. Econometrica 57, 1121–1152.
Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881–898.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65–91.
Jegadeesh, N., Titman, S., 1995. Short-horizon return reversals and the bid-ask spread. Journal of Financial Intermediation 4, 116–132.
Lewellen, J., 2011. The Cross-section of Expected Stock Returns. Manuscript. Dartmouth College, Hanover, NH.
Loughran, T., 1997. Book-to-market across firm size, exchange, and seasonality: is there an effect? Journal of Financial and Quantitative Analysis 32, 249–268.