Leverage and Performance Metrics in Asset Pricing

Alex R. Horenstein
Department of Economics
School of Business Administration
University of Miami
horenstein@bus.miami.edu

October 15, 2018

Abstract. Commonly used performance metrics calculated from tradable long-short strategies constructed using leverage lead to specious conclusions of statistical significance. In particular, even if assets are randomly assigned to the levered strategies' long and short portfolios, too often they generate high Sharpe Ratios and statistically significant CAPM, Carhart, and Fama-French Five Factor alphas. I call this finding *the mechanical effect of leverage on performance metrics* and suggest various tests to control for this effect when testing levered strategies.

Keywords. factor models, leverage, Sharpe Ratio, alpha, betting against beta, betting against alpha.

JEL classification. G10, G12.

1 Introduction

Since the development of the Intertemporal CAPM (Merton 1972) and the Arbitrage Pricing Theory (APT, Ross 1977), which allow for multifactor asset pricing models, the quest for risk factors other than the CAPM's market factor that can explain the cross-section of stock returns has been pervasive in the finance literature. Harvey et al. (2016) categorize 314 factors from 311 different papers published in top-tier finance journals and working papers between 1967 and 2014. The most common empirical method to create a tradable factor relies on finding a long-short strategy based on stock portfolios sorted, double sorted, and more recently triple sorted on some metrics (usually some firm characteristics, for example market capitalization and book to market value) that are not priced by other existing factors. A basic stylized fact shown in many papers to support a new variable as a factor is that portfolios sorted on certain metrics show a decreasing or increasing pattern in average excess returns. The most famous example is probably the three-factor model of Fama and French (1992), which augments the empirical CAPM with a factor capturing the size effect (SMB) and a factor capturing the value-premium (HML). Portfolios sorted on market capitalization show a decreasing pattern in average returns, and portfolios sorted on book to market value show an increasing pattern in average returns. Once these patterns are found, then some performance related metrics are calculated for the zero-net investment and dollar-neutral long-short strategy to confirm the finding of a risk factor. The most common performance metrics analyzed in the literature are the risk-premium of the factor, its Sharpe Ratio, and the magnitude and significance level of the alpha (pricing error) produced by regressing the proposed factor against some popular benchmark model.[1]

The added factors to the CAPM were needed because the empirical market portfolio (henceforth Market), usually the value-weighted portfolio's excess returns over the one-month T-bill, failed to explain the cross-section of stock returns. The Market beta-return relationship known as Security Market Line (SML) was found to be too flat.[2]

[1] Barillas and Shanken (2017) show that the test assets are generally not relevant when comparing asset pricing models with tradable factors. All that is needed is to check whether a proposed new factor is priced by the already existing ones.

[2] The Capital Asset Pricing Model (CAPM) developed by Treynor (1962), Sharpe (1964), Lintner (1965), and Mossin (1966) and extended by Black (1972) predicts a positive relationship between market betas and expected

Recently, Frazzini and Pedersen (FP, 2014) developed a compelling theoretical explanation for the observed flat SML. In their model, some investors are constrained in the leverage they can take, which leads them to tilt their portfolios toward high Market beta stocks to augment their expected returns. Consequently, high beta stocks are overpriced relative to low beta ones, which leads to a flat SML. The empirical implication of the FP model leads to a new type of factor constructed with leverage that they call Betting Against Beta (henceforth BAB). Their BAB factor is a market-neutral strategy with leverage that consists of buying a zero-net investment portfolio of low beta stocks' excess returns and selling a similar type of portfolio made of high beta stocks. Leverage is applied by using certain weights to magnify the excess return of the long portfolio and diminish the excess return of the short one. Horenstein (2018) follows FP's rationale to construct a strategy with leverage applied to assets sorted on the CAPM alpha. He calls it Betting Against Alpha (henceforth BAA).

Therefore, there seems to be room for a new paradigm to empirically find risk factors. If portfolios sorted on a certain metric or firm characteristic produce decreasing/increasing returns, a long-short strategy without leverage is a candidate to produce a new risk factor. If portfolios sorted on a certain metric or firm characteristic generate flat returns, then a long-short strategy with leverage is also a candidate to produce a new risk factor. In fact, as long as the long and short portfolios of a strategy produce a positive risk premium, leverage can produce positive risk premiums, whether these portfolios show decreasing, flat, or increasing patterns in excess returns. Therefore, having reliable metrics to assess the relevance of a strategy is a must to avoid the proliferation of factors that might not be relevant.

Focusing on the case of levered strategies, the goal of this paper is twofold. *First*, I investigate if applying leverage as in FP to a long-short strategy has a non-negligible impact on Sharpe Ratios and estimated alphas. I focus on Sharpe Ratios and alphas since these two metrics are arguably the most common metrics used in the empirical asset pricing literature to measure the performance of a tradable factor. *Second*, I investigate the performance metrics of several strategies constructed

returns. A puzzling finding from many empirical studies using linear factor models is that the empirical market portfolio has strong explanatory power for the comovement of stock returns, but its betas generate a flat SML. This result has been interpreted as a failure of the CAPM to explain the cross-sectional variation in expected stocks returns (e.g., Reinganum 1981, Lakonishok and Shapiro 1986, Fama and French 1992, and Fama and French 2004).

using leverage. The Market is not the only risk factor studied in the empirical finance literature. There are other traded and non-traded factors that could also be used to create levered strategies by betting against their factor loadings. Then, I investigate which of all the different levered strategies I create capture relevant information and how to use Sharpe Ratios and alphas to select among the candidate levered strategies.

To answer the first question, I start by simulating 100,000 strategies with leverage where I randomly assign assets to the low and high beta portfolios each year. I magnify the excess returns of the long (low beta) portfolio and diminish the excess returns of the short (high beta) portfolio using the same weights from FP's BAB strategy. From the 100,000 random strategies, I obtain an empirical distribution of their risk premiums; Sharpe Ratios; and CAPM, Carhart (1997), and Fama-French five factors model (FF5, 2015) alphas' t-statistics. I find that the randomly created long-short levered strategies generate positive risk premiums, high Sharpe Ratios, and alphas with t-statistics surpassing the hurdle of 3.0 suggested by Harvey et al. (2016) too often. In fact, the value 3.0 is included in the 95 percent interval of the alpha t-statistics' empirical distributions generated by the Carhart and FF5 models. I call this phenomena the *mechanical effect of leverage on performance metrics* (henceforth "the mechanical effect of leverage").[3] I also find in my simulations that using levered strategies like BAA and BAB to control for risk restores the relevance of the alphas' t-statistics, although not perfectly.

To answer the second question, I analyze sixteen strategies constructed using leverage, in which the low and high portfolios are sorted on the CAPM alpha (BAA), Market beta (BAB), the betas corresponding to the four factors augmenting the CAPM in FF5 (SMB, HML, RMW, and CMA), and the betas corresponding to ten macroeconomic variables.

To compare across strategies, I develop a benchmark scenario in which I use the same weights for the long and short portfolios in every levered strategy. I use as common weights those of the BAB strategy. Of course, keeping the low and high portfolios' weights constant across levered strategies

[3]In the FP model, the time-varying weights used to apply leverage in the BAB strategy are related to funding liquidity conditions (see FP's Proposition 4). Therefore, what I call the mechanical effect of leverage could be related to economic fundamentals. However, what is important in this study is that, using the same weights in the long-short portfolios across different levered strategies, I will be able to assess the impact of leverage on Sharpe Ratios and alphas, independently of whether the variable used to sort the strategies' portfolios is related to risk or not. This is why I call the average effect of leverage across strategies "mechanical."

might not lead to the optimal outcome of individual strategies in terms of high Sharpe Ratios or large expected risk premiums. However, my goal is not to find the weights that maximize the Sharpe Ratio or the risk premium of a particular levered strategy, but to provide guidance on how to use Sharpe Ratios and estimated alphas to assess the relevance of a strategy. Consistent with my simulation exercise, I find that most of the sixteen strategies produce large Sharpe Ratios and statistically significant alphas at less than the 1% level of significance.

Following Shanken and Barillas (2017), I analyze whether one levered strategy can be priced by another one in order to find the minimum set of levered strategies that can price all others. I find that the BAB strategy is not priced by any other levered strategies. On the other hand, the BAB strategy fails to price the BAA and some Macroeconomic BAB strategies. Importantly, the BAB and BAA strategies combined can price all other levered strategies. This is consistent with the results in Horenstein (2018), where he argues that BAB and BAA should subsume the information in all betting against beta strategies.

Given the previous results, I suggest the following three diagnostic tests using Sharpe Ratios and alphas for finding relevant strategies constructed with leverage: First, a strategy's Sharpe Ratio should be large enough to rule out being a product of the mechanical effect of leverage (at least 50% larger than the Market's Sharpe Ratio). Second, Sharpe Ratios should decrease across portfolios sorted on the metric used to construct the "betting against" levered strategy. Third, a new levered strategy should produce statistically significant alphas when risk is controlled by other levered strategies like BAB and BAA. Assessing the relevance of a levered strategy using alphas generated by models containing only factors constructed without leverage leads to false positives too often.

Of course there are other methods for testing whether a new proposed factor captures relevant information missed by the already existing ones. For example, Ahn et al. (2018) and Burnside (2016) suggest using rank estimation methods. Feng et al. (2017) propose a model selection method based on applying a double-selection LASSO in tandem with a two-pass regression. In this paper I focus on Sharpe Ratios and estimated alphas, arguably the two most common performance metrics used in the literature when analyzing tradable factors.

In a related study, Cederburg and O'Doherty (2016) explore the relevance of the BAB strategy and find that the alpha generated by it disappears once the estimation bias generated by the unconditional CAPM is controlled for in the conditional version of the model. They focus on the economic implications originated by the negative correlation between the portfolios sorted on betas and the equity premium. Liu et al. (2018) find that the low-beta anomaly disappears when they control for the relationship between beta and idiosyncratic volatility. Differently, in this paper I focus exclusively on the impact of leverage independently of the metric used to sort assets within portfolios. My results are relevant for studies about the low-beta anomaly as well as for any other strategy constructed using leverage.

To sum up, this paper adds to the literature on empirical asset pricing by studying the impact of leverage on Sharpe Ratios and estimated alphas' t-statistics, and by suggesting simple solutions to keep these metrics relevant for comparison purposes when analyzing levered strategies.

The rest of the paper is organized as follows. Section 2 describes the data I will use and explains the construction of the levered strategies. I also provide some preliminary statistics. Section 3 provides the quantitative analyses. I conclude in Section 4.

2 Data and construction of the levered strategies

2.1 Data and strategies

I use monthly data on US individual stock returns from the Center for Research in Security Prices (CRSP) from January 1968 until December 2015. The returns include dividends from common stocks traded on the NYSE, NASDAQ, and AMEX, excluding REITs and ADRs. Data on the Fama-French five factors model (FF5, 2015) as well as the Momentum factor for the Carhart (1997) model are from Kenneth French's website.[4] The FF5 factors are the excess market return (Market: the return on the CRSP value-weighted portfolio minus the return on the 1-month Treasury bill), Small Minus Big (SMB), High Minus Low (HML), Robust Minus Weak (RMW), and Conservative Minus Aggressive (CMA) factors. The Carhart factors are the Market, SMB, and HML factors from FF5 plus the Momentum factor (MOM, selling losers and buying winners 6 to 12 months ago). I also use data

[4] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

on macroeconomic variables from the Federal Reserve Bank of St. Louis.[5] Table 1 below shows the ten macroeconomic variables used in my analysis.

[Insert Table 1 around here]

A Betting Against Beta (Alpha) strategy consists of selling a portfolio containing high beta (alpha) stocks' excess returns over the risk free asset and buying a portfolio of low beta (alpha) stocks' excess returns, where the excess returns of the low beta (alpha) portfolio are amplified while the excess returns of the high beta (alpha) portfolio are reduced. I will analyze the effect of leverage on portfolios sorted using the following sixteen different variables:

[Insert Table 2 around here]

2.2 Construction of the levered factors

The Betting Against Beta (BAB) factor developed by Frazzini and Pedersen (FP, 2014) consists of selling a portfolio made of high Market beta stocks' excess returns over the risk free rate and buying a portfolio made of low Market beta stocks' excess returns (note that the long and short portfolios are both zero-net investments). Additionally, both portfolios are scaled by the inverse of the risky assets' weighted betas, and thus, given that the average Market beta value fluctuates around one, the excess returns of the low beta portfolio are amplified, while the excess returns of the high beta portfolio are reduced. As shown in FP, the BAB factor is a zero-net investment and market-neutral strategy, but it is not dollar-neutral in terms of the risky portfolios.

Modifying the weights used to augment and reduce the long-short portfolios' excess returns modifies the performance of the strategy. To simplify the comparison across different levered strategies, I will use the same weights on all of them. As shown in FP, further calibration of the weights applied to the long-short portfolio strategy can increase the resulting Sharpe Ratios. However, the goal in

[5] https://fred.stlouisfed.org

this paper is not to find the strategy with the largest Sharpe Ratio, but to compare the different possible strategies and identify those that are not priced by the already existing factors. For this purpose I create the following benchmark scenario. Every December I estimate the individual assets' betas and alphas using simple OLS regressions and five-year of monthly data. FP suggests calculating the parameters differently than the usual OLS, but again, my goal is not to find the highest Sharpe Ratio nor the largest pricing error but to compare across strategies. Therefore, using the same (most common) technique across strategies to calculate the parameters of interest seems appropriate in this case.[6]

Based on the estimated betas and alphas I create a low portfolio and a high portfolio, where the low portfolio contains the assets with estimated parameters below the median value and the high portfolio contains the assets with estimated parameters greater than or equal to the median. Then, the levered strategy consists of buying the low portfolio and selling the high portfolio with certain weights. It is important to stress again that I will use the same weights across all strategies. I will construct the weights as in Horenstein (2018), which closely follows the methodology implemented by FP. The next paragraphs explain how.

For each asset i, $\beta_i = \rho_{iM}\,(\sigma_i/\sigma_M)$, where ρ_{iM} is the correlation between the i asset's returns and the Market returns, while σ_i and σ_M are the asset i and Market estimated volatilities, respectively. As in FP, the ρ_{iM} is estimated using five year data while σ_i and σ_M are estimated using yearly data. The final Market beta assigned to an asset i (β_i^M) is compressed towards one as in FP using the formula $\beta_i^M = 0.6\beta_i + 0.4$.

Assets with β_i^M lower (higher) than the median are assigned to the low (high) beta portfolio and weighted using the same formula as in FP. More precisely, let nl be the number of assets in the low beta portfolio and zl be the $nl \times 1$ vector of beta ranks such that $zl_i = rank(\beta_i^M)$. The weight of an asset i in the low beta portfolio is given by $wl_i = (nl - zl_i + 1)/\sum zl_i$. Similarly, let nh be the number of assets in the high beta portfolio and zh be the $nh \times 1$ vector of beta ranks in this portfolio, where $zh_i = rank(\beta_i^M)$. The weight of an asset i in the high beta portfolio is given by $wh_i = zh_i/\sum zh_i$. Note that $\sum wl_i = \sum wh_i = 1$. The final weighted Market betas of the low and

[6]For example, the following papers also use 5 year of monthly data to calculate the parameters alpha and beta: Black et al. 1972, Banz 1981, Fama-French 1992, 1993, and 2015, just to mention some.

high beta portfolios are $\beta^L = \sum wl_i \beta_i^M$ and $\beta^H = \sum wh_i \beta_i^M$, respectively.

The main difference between this paper's calculation of the weights β^L and β^H and the one used in FP is that they use daily data to estimate betas and I use monthly data. Additionally, they allow for assets to have at least three years of data in their calculation, while I use assets with five years of data. On average, this paper's weights imply an investment of around \$1.67 in the long portfolio and \$0.63 in the short portfolio. Finally, they re-balance their strategies (and portfolio weights) monthly while I do it yearly (end of December).

The rationale that supports the modifications in the way I calculate the weights is related to my intention of creating a benchmark scenario to analyze the different levered strategies and can be summarized as follows. First, one of my goals is to compare across strategies using the same method (and weights) for every strategy. Second, the other goal of this paper is to check if some of the levered strategies cannot be priced by other factors, like for example the Fama-French five factors. Since the portfolios used to construct the FF factors are re-balanced annually, I decided to use the same frequency to re-balance the strategies in this paper to compare homogeneous scenarios. Strategies with different re-balancing frequencies generate different transaction costs, increasing the likelihood of data-snooping problems (Novy-Marx and Velikov, 2016). Finally, as it is standard in the literature, I use 60 months of data (5 years) for estimating alphas and betas. Importantly, Horenstein (2018) shows that his methodology for calculating the weights β^L and β^H does not substantially differ from that of FP.

Now I summarize relevant details necessary to construct the different levered strategies used in this paper. I obtain the assets' betas by simple OLS regression. Macroeconomic variables are used with a 6-month lag to be sure that revised series are available at the moment in which the econometrician runs the regression. Portfolio re-balancing is done each December and the assets in the low and high portfolios are fixed for twelve months. Alphas used for the BAA strategy are also calculated by OLS. The returns of the low and high portfolio are $r^L = \sum wl_i r_i^L$ and $r^H = \sum wh_i r_i^H$, respectively. The BAB factor's monthly returns for a given year, with yearly rebalancing at the end of December, are $r_{t+s}^{BAB} = \frac{1}{\beta_t^L}(r_{t+s}^L - r_{t+s}^f) - \frac{1}{\beta_t^H}(r_{t+s}^H - r_{t+s}^f)$, where $s = 1, ..., 12$ and t corresponds to December. The same weights β^L and β^H apply to all strategies.

Figure 1 below shows the monthly Sharpe Ratios generated by all of the strategies analyzed in this paper over the monthly Sharpe Ratio of the Market factor.

[Insert Figure 1 around here]

The Sharpe Ratio of the BAB strategy is the largest and more than doubles the Sharpe Ratio of the Market factor. Interestingly, all strategies except the HML BAB and CMA BAB have Sharpe Ratios larger than the Market factor. BAA has a large Sharpe Ratio too, almost as large as the one of BAB. Similarly, several strategies based on macroeconomic variables present sizable Sharpe Ratios.

In the next Section I will show that these large Sharpe Ratios might be misleading. More precisely, if assets are sorted on a variable unrelated to risk generating a flat SML, a levered strategy might still produce large Sharpe Ratios relative to the Market factor simply because of what I call the mechanical effect of leverage.

3 Results

3.1 The mechanical effect of leverage

A variable used as a factor but unrelated to risk will most likely lead to a flat SML when sorting portfolios on the estimated factor loadings. Let's call the factor loadings generated by a variable unrelated to risk *random beta*, to distinguish it from the standard (risk-related) beta. However, by chance it can happen that ranking assets by some random beta assigns a higher proportion of low risk-related beta stocks to the low portfolio than to the high portfolio, making an irrelevant variable appear relevant when used to create a levered strategy. Therefore, having an empirical distribution of Sharpe Ratios and alpha's t-statistics over several betting against beta strategies based on random betas will help us understand the impact of using leverage on these performance metrics.

In this Section I generate the distribution of Sharpe Ratios and alpha's t-statistics of 100,000

randomly created portfolios to study the mechanical effect of leverage using a simple bootstrap.[7] Using US stock return data, every December I randomly assign half of the stocks to a "low" portfolio and the other half to a "high" portfolio. Then, I randomly assign each assets' i rank zl_i (zh_i) in the low (high) portfolio. Remember that the weight of an asset in a portfolio depends on its rank (see Section 2.2).

As with the strategies described in Table 2, I rebalance the long and short random portfolios once a year, at the end of December. I do this for data from December 1972 to December 2014, which means I generate monthly data on these random portfolios from January 1973 to December 2015. Then, I apply leverage to the long and short random portfolios using the weights β_t^L and β_t^H from FP's BAB factor described in the previous Section. With the leverage applied to the portfolios I create the long-short levered random strategy and calculate its monthly Sharpe Ratio, CAPM alpha t-statistics, Carhart alpha t-statistics, and FF5 alpha t-statistics.[8] Additionally, I also compute the Sharpe Ratios and alpha t-statistics of the unlevered low portfolio excess returns, unlevered high portfolio excess returns, and the unlevered low minus high portfolio strategy (unlevered random strategy) for comparison purposes.[9] I repeat this process 100,000 times to get an empirical distribution of the (monthly) Sharpe Ratios and alpha t-statistics. The only parameters that remain constant in each iteration of the bootstrap are the weights β_t^L and β_t^H assigned to the low and high portfolios to generate leverage.

Figure 2 shows the distribution of the parameters of interest for the levered random strategies.

[Insert Figure 2 around here]

A conservative interpretation of the figure is that we cannot rule out that the only effect a levered

[7] To calculate the t-statistics I use heteroskedastic robust standard-errors.

[8] To avoid aggregation issues, I do not annualize the estimated monthly Sharpe Ratios (see Lo 2002).

[9] The unlevered random strategy is simply the result of subtracting the high portfolio excess returns from the low one without using leverage. In terms of the discussion in Section 2.2, the levered strategy's returns are $r_{t+s}^{BAB} = \frac{1}{\beta_t^L}(r_{t+s}^L - r_{t+s}^f) - \frac{1}{\beta_t^H}(r_{t+s}^H - r_{t+s}^f)$ while the unlevered ones are simply $r_{t+s}^{UnBAB} = r_{t+s}^L - r_{t+s}^H$, where $s = 1, ..., 12$ and t corresponds to December.

strategy is capturing is the mechanical effect of leverage when the monthly Sharpe Ratios lie in the range (0.13, 0.16). The average value of the Sharpe Ratio's distribution is around 0.15. For the (unlevered) FF5 Market, SMB, HML, RMW, and CMA factors, the corresponding Sharpe Ratios are 0.11, 0.07, 0.12, 0.11, and 0.17. Therefore, the mechanical effect of leverage produces on average higher Sharpe Ratios than any FF5 factor but CMA – up to 50% larger than the Market factor. This result shows the importance of taking into account this effect when analyzing levered strategies in order to avoid false positives.

Additionally, random levered strategies produce on average sizable alpha t-statistics (for the Carhart and FF5 models the average alpha t-statistic is around 3). Therefore, this commonly used metric to analyze a long-short strategy leads to false positives too often and does not seem appropriate to assess the performance of levered strategies without further corrections.

To further study the mechanical effect of leverage, I will compare the distribution of the Sharpe Ratios and alphas' t-statistics of the simulated levered random strategies with those of the simulated unlevered random strategies.

The summary statistics are in Table 3, the histograms comparing the parameter's distribution from the levered random strategies against the unlevered ones are in Figure 3, and the histograms comparing the parameter's distribution from the levered random strategy against the random Low Portfolio are in Figure 4.[10]

[Insert Table 3 around here]

[Insert Figure 3 around here]

[Insert Figure 4 around here]

First let's compare Sharpe Ratios. From Panel (b) of Table 3 and Panel (a) of Figure 3 we observe that the Sharpe Ratios of the unlevered random strategies have a mean of zero. This is expected since the average values of this parameter for the Low and High portfolio are the same as

[10] I do not show the histograms comparing the parameter's distribution from the simulated levered random strategies to that from the simulated random High portfolios since, as expected, they are almost identical to those of the simulated random Low portfolio.

shown in Panels (c) and (d) of Table 3. However, the average monthly Sharpe Ratio of the levered random strategy is 0.147, a non-negligible 31.5% larger than the Market portfolio's monthly Sharpe Ratio during the same period of analysis (0.111).

The mean and standard deviation of the levered random strategies' Sharpe Ratio are also larger than those of the Low and High portfolios separately. The average Sharpe Ratio is larger for the random levered strategy with respect to the low portfolio excess return by 9% (and the difference is significant at less than 1%), while the maximum value of the Sharpe Ratio for the random strategy is larger by 15%. Overall, I find that this particular leverage increases the Sharpe Ratio by an average of 9% even if the beta used to construct the portfolio is unrelated to risk; therefore, the mechanical effect of leverage magnifies the Sharpe Ratio of the levered strategies.

Now let's analyze the CAPM, Carhart and FF5 alphas' t-statistics. The unlevered random strategies produce on average alphas' t-statistics equal to zero [Table 3 Panel (b)]. Panels (c) and (d) of Table 3 show that the CAPM and FF5 alphas' t-statistics are on average significantly less than 2 for the Low and High portfolio, while the average Carhart alpha's t-statistic is higher than 2.5. However, a t-statistic of 3 is not in the 95% interval of any model's t-statistics empirical distribution for the portfolios' case. Therefore, the threshold suggested by Harvey et al. (2016) of using t-statistics greater than 3 seems reasonable for unlevered portfolios or strategies. Once we apply leverage to the random portfolios the alphas' t-statistics generated by Carhart and FF5 increase dramatically, as depicted by an average t-statistic of around 3 for both models. In this case, the mechanical effect of leverage leads to false positives too often when testing the significance of alpha. Surprisingly, for the case of levered strategies the CAPM seems to generate statistically significant alphas less often than Carhart and FF5, but its simulated t-statistics are still larger than 2 on average.

From the previous results it is clear that alpha is not an informative performance metric for levered strategies when common risk is controlled using unlevered factors. Can the value of alpha as a performance metric be restored if risk is controlled using levered strategies? To study this possibility I now run the same 100,000 simulations as before. However, now I estimate the alphas of the simulated random levered strategies using BAB and BAA, separately and together, as independent variables.

I also control for common risk augmenting the CAPM, Carhart, and FF5 models with BAB and BAA separately. Overall, I use 9 models to control for common risk. Table 4 shows the different alpha t-statistics results from this new simulation exercise.

[Insert Table 4 around here]

From Table 4 we can observe that controlling with just BAB seems to make the alpha t-statistic informative again. Although the average t-statistic from the 100,000 simulations is still not centered at zero, the value of 2 is not in the range of the estimated alphas' t-statistics. As a conservative measure, selecting models producing a maximum t-statistic with an absolute value of 2 in the 100,000 iterations, the following models seem to make the alpha t-statitistic informative when testing strategies with leverage: BAB, BAA, CAPM+BAB, and Carhart+BAA. Interestingly, models with more factors than just the Market sometimes reduce the benefits of adding BAB as a factor while the opposite happens for BAA. In fact, the only model that seems to generate a distribution of t-statistics centered close to zero is Carhart+BAA.

To sum up the results from the simulations, I found that the impact of the mechanical effect of leverage on both parameters, Sharpe Ratios and alphas' t-statistics, is non-negligible. Once I apply leverage to a long-short strategy, the possibilities of confounding an irrelevant factor as a relevant one are quite high. The simulations also show that controlling for risk using levered strategies helps reduce the probability of having a false positive. A minimum test that a levered strategy must pass if alpha is used as a performance metric is that it should generate a statistically significant alpha when controlling for BAB. If instead the researcher wants to control for risk using BAA, Carhart+BAA seems to provide reliable results.

In the next Section, I analyze more in depth the performance metrics of the strategies described in Table 2, as well as the possibility that some of them appear relevant simply because of the mechanical effect of leverage.

3.2 Portfolio characteristics and performance of strategies with leverage

In this Section I analyze the levered strategies presented in Table 2. I start by showing in Table 5 the summary statistics for the sixteen levered strategies between January 1973 and December 2015.[11] For each strategy the table reports summary statistics for the low portfolio, high portfolio, and levered strategy (levered low portfolio minus delevered high portfolio). It reports the monthly average excess return, monthly CAPM alpha, monthly Carhart alpha, monthly FF5 alpha, monthly Sharpe Ratio, and average size of the weighted portfolios (in thousands of USD) at the time of re-balancing.

[Insert Table 5 around here]

The table shows that almost all levered strategies produce CAPM, Carhart, and FF5 alphas that are statistically significant at the 1% level or less. Additionally, I already showed in Figure 1 that all strategies except two produce Sharpe Ratios that are higher than that of the Market factor. Therefore, without further analysis, almost all levered strategies seem worth pursuing.

In the previous Section we find that levered strategies with monthly Sharpe Ratios below 0.16 might not be relevant. This allows us to remove from the set of relevant strategies the RMW BAB strategy and many of the strategies based on macroeconomic variables (ΔLabor Participation, ΔPersonal Savings Rate, ΔReal Personal Consumption, and ΔNew One Family Houses Sold). However, this metric still keeps as possibly relevant strategies 9 out of the 16 strategies. Therefore, if a researcher choses relevant strategies based solely on these two metrics – alphas and Sharpe Ratios – without further analysis, then still too many strategies seem worth pursuing.

Panels A through F report the results for the levered strategies using the CAPM alpha and the betas from each of the FF5 factors individually. The highest Sharpe Ratio is that of the BAB strategy, followed by BAA and the SMB BAB strategy. The main difference between the BAB, BAA, and SMB BAB strategies with respect to the others is that they show a decreasing Sharpe Ratio on the unlevered portfolios used to construct the strategies. Therefore, a decreasing Sharpe

[11] Our stock return data starts in January 1968 (macroeconomic data starts in July 1967), but our levered strategies data start in January 1973 since I use five years of data to initialize the calculations.

Ratio between the long and short unlevered portfolios is important for the performance of a strategy constructed using leverage.

Panels G to P show results from strategies formed on sensitivities to macroeconomic variables. Some of the strategies present decreasing Sharpe Ratios on the sorted portfolios while others do not. However, all of them present statistically significant alphas when regressing the strategies against the CAPM, Carhart, and FF5 factor models. These results warn us again about some levered strategies appearing successful simply because of the leverage (and deleverage) applied to portfolios with similar expected returns and/or Sharpe Ratios, independently of whether the betas used to sort the assets are related to a source of common risk or not. Nevertheless, there are still some macroeconomic variables that generate sizable monthly Sharpe Ratios for their levered strategies, while showing decreasing Sharpe Ratios in the sorted portfolios, like the Default Spread BAB strategy shown in Panel P or the ΔUnemployment BAB strategy shown in Panel H.

Summarizing the results in this section, first we can conclude that the levered strategies that present the highest Sharpe Ratios are those whose sorted portfolios on betas show a decreasing pattern in Sharpe Ratios. Second, a decreasing pattern in Sharpe Ratios is more important than a decreasing pattern in expected returns to generate levered strategies with a high Sharpe Ratio. The strategy with the highest Sharpe Ratio is BAB, but the portfolios sorted by Market beta show an increasing pattern in average excess returns, while many other portfolios show a flat or even decreasing pattern of excess returns (e.g. HML BAB in Panel D or ΔPersonal Savings Rate in Panel J).

Although the mechanical effect of leverage can lead to high Sharpe Ratios and statistically significant abnormal returns independently of whether the factor loadings are related to a risk or not, many of the strategies I analyzed in this Section present Sharpe Ratios that are too high to be considered the sole product of the mechanical effect of leverage. These strategies are BAA, BAB, SMB BAB, Inflation BAB, ΔUnemployment BAB, ΔAvg Hour Earn BAB, ΔNew Housing Auth BAB, ΔNew Houses Prc BAB, and Default Spread BAB. Now the obvious question is how many of these strategies are redundant. I study this question in the next Section.

3.3 The minimum set of levered strategies that price all others

The results of the previous Section show that the high Sharpe Ratios of some of the levered strategies I analyzed are unlikely to be due to the mechanical effect of leverage (especially those whose Sharpe Ratios are around 0.20 or more). An important question that remains to be answered, especially for those interested in constructing better asset pricing models, is how many of these levered strategies are not priced by the others. To answer this I use insight from Barillas and Shanken (2017). They showed that for the case of tradable factors "it turns out that test assets tell us nothing about model comparison, beyond what we learn by examining the extent to which each model prices the factors in the other models." In other words, to compare factor models we can simply regress one factor against another set of factors and see if the pricing error is statistically significant. If it is, then the factor used as an independent variable can be added to the model to improve it. Importantly, the results in Section 3.1 show that regressing levered strategies on models with only unlevered factors produces statistically significant alphas too often. The simulations in Section 3.1 also show that the relevance of alpha can be restored once levered strategies are used to control for common risk. This Section's results will further show that levered strategies' alphas become insignificant once we control for a model containing levered factors.

I first run univariate regressions of all strategies against each other to study which ones can price other strategies and which ones cannot. Table 6 below shows the significance levels of 240 regressions obtained from regressing each of the sixteen levered strategies on the other fifteen. Rows are assigned to the strategies used as dependent variables while columns are assigned to the strategies used as regressors. For example, the first row corresponds to the significance level of the alpha t-statistics obtained using the BAB strategy as dependent variable and all other strategies as independent. We see from the first row that the BAB strategy is not priced by any other strategy, since the pricing errors of the univariate regressions are statistically significant at the 1% level or less, as shown by the three stars in every cell.

[Insert Table 6 around here]

Additionally, the first column shows that the BAB strategy fails to price BAA and two Macroeconomic BAB strategies: ΔUnemployment BAB and Default Spread BAB. The two levered macroeconomic strategies that BAB fails to price are priced by BAA, as shown by the column corresponding to BAA. Therefore, it seems that BAB and BAA contain all the relevant information for pricing purposes of these sixteen strategies.

As a final analysis, I test if the models I found relevant for pricing levered strategies in the simulation exercise can price the levered strategies analyzed in this section. Table 7 shows the value of the pricing errors of the strategies and their level of significance when regressed against the BAB, CAPM+BAB, and Carhart+BAA models. I add a last column using BAA+BAB to corroborate the conclusion in the previous paragraph that a model containing these two strategies suffices to price all the others.

[Insert Table 7 around here]

The first column of the table adds numerical values to what I already showed with stars in Table 6: the BAB factor fails to price ΔUnemployment BAB, and Default Spread BAB. Therefore, these two macroeconomic levered strategies (in addition to BAA) seem to contain relevant information missed by BAB. The Carhart+BAA model seem to price all strategies except SMB BAB and Default Spread BAB. When using this model most alphas are insignificant, which is consistent with the simulations' results. Finally, as discussed in the previous paragraph, the last column of the table shows that a model with BAA and BAB price all the other fourteen levered strategies.

To sum up, the results show that if a researcher wants to use alpha as a metric for assessing the relevance of a factor constructed using leverage, controlling for common risk using a model with levered factors is a must. The usual models with unlevered factors, for example Carhart or FF5, do not seem reliable for testing levered strategies.

4 Concluding remarks

This paper is concerned with the use of Sharpe Ratios and alphas' t-statistics when assessing long-short strategies constructed using leverage like Frazzini and Pedersen's (2014) Betting Against Beta (BAB) and Horenstein's (2018) Betting Against Alpha (BAA).

Simulating 100,000 strategies that randomly assign assets to the low and high portfolio of a levered strategy, and using similar weights to apply leverage as in Frazzini and Pedersen's BAB strategy, I find that monthly Sharpe Ratios are on average 20% larger – and up to 50% larger – than that of the empirical Market factor. These simulated strategies also tend to produce CAPM, Carhart, and FF5 alpha t-statistics that are highly significant, easily surpassing the hurdle of 3.0 suggested by Harvey et al. (2016). I call this the mechanical effect of leverage on performance metrics.

For testing purposes I also create 16 different levered strategies using the same weights that apply leverage to the BAB strategy. To assign assets to the low and high portfolios, 15 of these 16 strategies sort the stocks by betas generated by the FF5 factors individually and ten macroeconomic variables. In addition, the BAA strategy assigns assets to the low and high portfolio based on the CAPM's alpha. I find that many of these strategies seem useful and produce sizable Sharpe Ratios that cannot be explained by the mechanical effect of leverage. However, and consistent with the results in Horenstein (2018), I find that two levered strategies suffice to price all others: BAB and BAA.

Based on all these findings I suggest three simple diagnostic tests for checking the relevance of a tradable long-short strategy constructed using leverage. First, a strategy's monthly Sharpe Ratios should be large enough to rule out being a product of the mechanical effect of leverage (at least 50% larger than the Market factor's Sharpe Ratio). Second, Sharpe Ratios should decrease across portfolios sorted on the metric used to construct the levered strategy. Third, a new levered strategy should produce statistically significant alphas when the risk is controlled by a model containing levered factors. Controlling common risk using models with only unlevered factors like the standard CAPM, Carhart, or FF5 models to estimate the pricing errors of a levered strategy produces false positives too often.

Of course there are other methods for testing whether a new proposed factor captures relevant information missed by the already existing ones. For example, Ahn et al. (2018) and Burnside (2016) suggest using rank estimation methods. Feng et al. (2017) propose a model selection method based on applying a double-selection LASSO in tandem with a two-pass regression. Horenstein (2018) uses rank estimation methods to confirm the relevance of the BAA and BAB strategies. In this paper I focus on Sharpe Ratios and estimated alphas, arguably the two most common performance metrics used in the literature when analyzing tradable factors.

References

[1] Ahn S.C., A. Horenstein, and N. Wang, 2018, Beta matrix and common factors in stock returns, *Journal of Financial and Quantitative Analysis* 53, 1417-1440.

[2] Banz R., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

[3] Barillas F. and J. Shanken, 2017, Which alpha?, *Review of Financial Studies* 30, 1316-1338.

[4] Black F., 1972, Capital market equilibrium with restricted borrowing, *Journal of Business* 45, 444-455.

[5] Black F., M. Jensen, and M. Scholes, 1972, The capital asset pricing model: Some empirical tests, *working paper available* https://ssrn.com/abstract=908569

[6] Burnside C., 2016, Identification and inference in linear stochastic discount factor models with excess returns, *Journal of Financial Econometrics* 14, 295–330.

[7] Carhart M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

[8] Cederburg S. and M. O'Doherty, 2016, Does it pay to bet against beta? On the conditional performance of the beta anomaly, *Journal of Finance* 71, 737–774.

[9] Fama E. and K. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

[10] Fama E. and K. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

[11] Fama E. and K. French, 2004, The Capital Asset Pricing Model: Theory and evidence, *Journal of Economic Perspectives* 18, 25-46.

[12] Fama E. and K. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1-22.

[13] Feng G., S. Giglio, and D. Xiu, 2017, Taming the factor zoo, *working paper*, available at SSRN: https://ssrn.com/abstract=2934020.

[14] Frazzini A. and L. Pedersen, 2014, Betting against beta, *Journal of Financial Economics*, 111, 1-25.

[15] Harvey C., Y. Liu, and H. Zhu, 2016, ... and the cross-section of expected returns, *Review of Financial Studies* 29, 5-68.

[16] Horenstein A., 2018, Can academic research generate new anomalies?, *working paper*, available at SSRN: https://ssrn.com/abstract=3054718.

[17] Lakonishok J. and A. Shapiro, 1986, Systematic risk, total risk and size as determinants of stock market returns, *Journal of Banking and Finance* 10, 115-132.

[18] Lintner J., 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economic and Statistics* 47, 13-37.

[19] Liu J., R. Stambaugh, and Y. Yuan, 2018, Absolving beta of volatility's effects, *Journal of Financial Economics* 128, 1-15.

[20] Lo A., 2002, The statistics of Sharpe ratios, *Financial Analysts Journal* 58, 36-52.

[21] Merton R., 1972, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

[22] Mossin J., 1966, Equilibrium in a capital asset market, *Econometrica* 35, 768-783.

[23] Novy-Marx R. and M. Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104-147.

[24] Reinganum M., 1981, A new empirical perspective on the CAPM, *Journal of Financial and Quantitative Analysis* 16, 439-462.

[25] Ross S., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

[26] Sharpe W., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

[27] Treynor J. Toward a theory of market value and risky assets. Unpublished manuscript. Final version in *Asset Pricing and Portfolio Performance*, 1999, Robert A. Korajczyk ed., London: Risk Books, 15-22.

Table 1: Macroeconomic variables

This table describes the ten macroeconomic variables used in our study. The first column reports the name of the variable. The second column explains how the variable was constructed, and the third column reports the name of the variable in the FRED database. All variables have been downloaded from the Federal Reserve Bank of St. Louis (https://fred.stlouisfed.org).

Variable	Construction	FRED Series
Inflation	Monthly change in the Consumer Price Index	CPIAUCSL
Δ Unemployment	Monthly change in Civilian Unemployment Rate	UNRATE
Δ Labor Part	Monthly change in Civilian Labor Force Participation Rate	CIVPART
Δ Pers Savings	Monthly change in Personal Saving Rate	PSAVERT
Δ Real Pers Cons	Monthly change in Real Personal Consumption Expenditures	DPCERAM1M225NBEA
Δ Avg Hour Earn	Monthly change in Average Hourly Earnings of Production and Nonsupervisory Employee	AHETPI
Δ New Housing Auth	Monthly change in New Private Housing Units Authorized by Building Permits	PERMIT
Δ New Houses Prc	Monthly change in Median Sales Price for New Houses Sold	MSPNHSUS
Δ New Houses Sold	Monthly change in New One Family Houses Sold	HSN1F
Default Premium	Difference between BAA-rated and AAA-rated corporate bonds	BAA10YM, AAA10YM

Table 2: Levered strategies' descriptions

This table describes the sixteen levered long-short strategies used in this study. Each strategy consists of selling a portfolio containing high beta (alpha) stocks' excess returns over the risk-free asset and buying a portfolio of low beta (alpha) stocks' excess returns. There are fifteen strategies based on sorting assets by betas obtained from univariate OLS regressions. Five of them use the betas from each of the Fama-French five factors (BAB, SMB BAB, HML BAB, RMW BAB, and CMA BAB) and ten use the betas from the macroeconomic variables described in Table 1. The macroeconomic variables are used with a 6-month lag to guarantee that revised series are available at the moment in which the econometrician runs the regression. Additionally, BAA is constructed by sorting assets on their CAPM alphas. Individual data on stock returns comes from the CRSP database, while the data for the CAPM, Carhart, and FF5 models comes from Kenneth French's webpage. The macroeconomic variables are from the Federal Reserve Bank of St. Louis webpage.

Strategy	Short description
BAA	Strategy short on assets with higher than the median CAPM alpha and long in assets with lower than the median CAPM alpha
BAB	Strategy short on assets with higher than the median MKT beta and long in assets with lower than the median MKT beta
SMB BAB	Strategy short on assets with higher than the median SMB beta and long in assets with lower than the median SMB beta
HML BAB	Strategy short on assets with higher than the median HML beta and long in assets with lower than the median HML beta
RMW BAB	Strategy short on assets with higher than the median RMW beta and long in assets with lower than the median RMW beta
CMA BAB	Strategy short on assets with higher than the median CMA beta and long in assets with lower than the median CMA beta
Inflation BAB	Strategy short on assets with higher than the median Inflation's loading and long in the others
Δ Unemployment BAB	Strategy short on assets with higher than the median Δ Unemplyment's loading and long in the others
Δ Labor Part BAB	Strategy short on assets with higher than the median Δ Labor Part's loading and long in the others
Δ Pers Savings BAB	Strategy short on assets with higher than the median Δ Pers Savings's loading and long in the others
Δ Real Pers Cons BAB	Strategy short on assets with higher than the median Δ Real Pers Cons's loading and long in the others
Δ Avg Hour Earn BAB	Strategy short on assets with higher than the median Δ Avg Hour Earn's loading and long in the others
Δ New Housing Auth BAB	Strategy short on assets with higher than the median Δ New Housing Auth's loading and long in the others
Δ New Houses Prc BAB	Strategy short on assets with higher than the median Δ New Houses Prc's loading and long in the others
Δ New Houses Sold BAB	Strategy short on assets with higher than the median Δ New Houses Sold's loading and long in the others
Default Premium BAB	Strategy short on assets with higher than the median Default Premium's loading and long in the others

Table 3: Simulations' summary statistics when controlling risk using only unlevered factors

This table presents the statistics of different performance metrics' distributions calculated using 100,000 simulations. The statistics are the mean, standard deviation, maximum, and minimum values of the generated empirical distribution. The performance metrics are the Sharpe Ratios and the heteroskedastic robust t-statistics generated by regressing each simulated strategy on the following three models: CAPM, Carhart, and FF5. The simulated strategies are generated by randomly assigning half of the stocks to a "low" portfolio and the other half to a "high" portfolio at the end of every December. Then, I randomly assigned each asset's i rank zl_i (zh_i) in the low (high) portfolio. The weight of an asset in a portfolio depends on its rank (see Section 2.2). I do this for data from December 1972 to December 2014, which means I generate monthly data on these random portfolios from January 1973 to December 2015. Panel (a) shows the distribution results from the 100,000 levered strategies (low-high portfolio with leverage), where I apply the same leverage to each strategy (see Section 2.2). Panel (b) shows the statistics for the performance metrics of a strategy consisting of subtracting the high portfolio from the low one without applying leverage. Panels (c) and (d) show the statistics for the performance metrics of the (unlevered) Low and High portfolio, respectively. Individual data on stock returns comes from the CRSP database, while the data for the CAPM, Carhart, and FF5 models comes from Kenneth French's webpage.

	Mean	Std. Dev.	Min	Max	95% Interval
(a) Levered Strategies					
Sharpe Ratio	0.147	0.004	0.131	0.163	[0.140 , 0.154]
CAPM's alpha t-statistic	2.195	0.143	1.554	2.840	[1.917 , 2.475]
Carhart's alpha t-statistic	2.940	0.172	2.140	3.734	[2.606 , 3.278]
FF5's alpha t-statistic	3.118	0.195	2.288	3.997	[2.739 , 3.503]
(b) Unlevered Strategies					
Sharpe Ratio	0.000	0.044	-0.204	0.214	[-0.086 , 0.086]
CAPM's alpha t-statistic	-0.001	1.005	-4.697	4.940	[-1.960 , 1.967]
Carhart's alpha t-statistic	-0.002	0.995	-4.443	4.787	[-1.957 , 1.959]
FF5's alpha t-statistic	-0.004	1.038	-4.717	4.903	[-2.040 , 2.046]
(c) Low Portfolio					
Sharpe Ratio	0.135	0.002	0.128	0.142	[0.132 , 0.138]
CAPM's alpha t-statistic	1.772	0.072	1.461	2.108	[1.631 , 1.913]
Carhart's alpha t-statistic	2.608	0.118	2.064	3.156	[2.377 , 2.840]
FF5's alpha t-statistic	1.101	0.112	0.622	1.622	[0.883 , 1.322]
(d) High Portfolio					
Sharpe Ratio	0.135	0.002	0.128	0.142	[0.132 , 0.138]
CAPM's alpha t-statistic	1.772	0.072	1.434	2.083	[1.631 , 1.912]
Carhart's alpha t-statistic	2.608	0.118	2.059	3.167	[2.377 , 2.840]
FF5's alpha t-statistic	1.102	0.112	0.582	1.588	[0.882 , 1.321]

Table 4: Simulations' summary statistics when controlling risk using levered factors

This table presents simulation results for t-statistics' empirical distributions calculated using 100,000 simulated levered strategies. The statistics are the mean, standard deviation, maximum, and minimum values of the generated empirical distribution of the heteroskedastic robust t-statistics generated by regressing each simulated strategy on the following nine models: BAB, BAA, BAB+BAA, CAPM+BAB, Carhart+BAB, FF5+BAB, CAPM+BAA, Carhart+BAA, and FF5+BAA. The simulated strategies are generated by randomly assigning half of the stocks to a "low" portfolio and the other half to a "high" portfolio at the end of every December. Then, I randomly assign each asset's *i* rank zl_i (zh_i) in the low (high) portfolio. The weight of an asset in a portfolio depends on its rank (see Section 2.2). I do this for data from December 1972 to December 2014, which means I generate monthly data on these random portfolios from January 1973 to December 2015. Betas and alphas used to assign the assets to the low and high portfolios of BAA and BAB are estimated with OLS. I use 5-year data to calculate betas and alphas. Portfolios are rebalanced yearly, at the end of December. I use monthly data from January 1968 – December 2015 to construct portfolios for the period January 1973 – December 2015. Individual data on stock returns comes from the CRSP database, while the data for the CAPM, Carhart, and FF5 models comes from Kenneth French's webpage.

	Mean	Std. Dev.	Min	Max	95% Interval
BAB alpha t-statistic	1.690	0.069	1.359	1.997	[1.554 , 1.825]
BAA alpha t-statistic	-1.584	0.155	-2.256	-0.925	[-1.887 , -1.280]
BAB+BAA alpha t-statistic	-1.710	0.153	-2.374	-1.072	[-2.010 , -1.413]
CAPM+BAB alpha t-statistic	1.192	0.125	0.582	1.742	[0.947 , 1.438]
Carhart+BAB alpha t-statistic	2.164	0.170	1.409	2.951	[1.831 , 2.501]
FF5+BAB alpha t-statistic	2.200	0.183	1.443	3.037	[1.843 , 2.561]
CAPM+BAA alpha t-statistic	-1.743	0.228	-2.714	-0.771	[-2.189 , -1.297]
Carhart+BAA alpha t-statistic	-0.058	0.240	-1.082	0.994	[-0.528 , 0.411]
FF5+BAA alpha t-statistic	1.092	0.287	-0.118	2.323	[0.536 , 1.662]

Table 5: Levered strategies' summary statistics

This table presents the monthly performance metrics for the low beta/alpha excess returns portfolios, high beta/alpha excess returns portfolios, and low minus high strategy with leverage using realized CAPM alpha (Panel A) and realized betas from univariate regressions on the Market (Panel B), SMB (Panel C), HML (Panel D), RMW (Panel E), and CMA (Panel F). Panels G to P use betas from univariate regression on macroeconomics variables to construct the low and high portfolios. The ten macroeconomic variables are described in Table 1. The macroeconomic variables are used with a 6-month lag to guarantee that revised series are available when the econometrician runs the regression. The performance metrics are the monthly Sharpe Ratios, monthly excess returns over the one-month T-bill, and abnormal returns for the CAPM, Carhart (1997), and Fama-French Five Factor (FF5, 2015) models. The abnormal returns are estimated by OLS, and their significance levels are calculated using heteroskedastic robust standard errors. The variable Size corresponds to the weighted market cap value of the portfolios. The Low, High, and Low-High portfolios are weighted with the same weights across strategies. The column Low (High) shows the results for the portfolio containing assets with realized betas/alphas below (above) the median value at the moment of rebalancing. Betas and alphas used to assign the assets to the low and high portfolios are estimated with OLS. I use 5-year data to calculate betas and alphas. Portfolios are rebalanced yearly, at the end of December. I use monthly data from January 1968 – December 2015 to construct portfolios for the period January 1973 – December 2015. Individual data on stock returns comes from the CRSP database, while the data for the CAPM, Carhart, and FF5 models comes from Kenneth French's webpage. The macroeconomic variables are from the Federal Reserve Bank of St. Louis webpage.

	Low Pflio	High Pflio	Low - High	Low Pflio	High Pflio	Low - High	Low Pflio	High Pflio	Low - High	Low Pflio	High Pflio	Low - High
		A) BAA			E) RMW BAB			I) Δ Labor Participation BAB			M) Δ New Private Housing BAB	
Excess Return	1.15%	0.67%	1.60%	1.01%	0.71%	1.33%	0.93%	0.88%	1.06%	0.95%	0.85%	1.12%
CAPM alpha	0.60%***	0.10%	1.10%***	0.38%**	0.25%**	0.58%*	0.36%***	0.34%***	0.48%***	0.39%***	0.30%**	0.56%***
Carhart alpha	0.55%***	0.06%	1.12%***	0.42%***	0.13%*	0.87%***	0.25%***	0.31%***	0.41%***	0.33%***	0.24%***	0.60%***
FF5 alpha	0.33%***	0.00%	0.90%***	0.37%***	-0.08%	1.07%***	0.19%***	0.12%*	0.55%***	0.23%***	0.08%	0.63%***
Sharpe Ratio	0.18	0.12	0.22	0.15	0.15	0.13	0.16	0.16	0.16	0.16	0.15	0.17
Size	$1,358,710	$1,990,035		$1,065,041	$2,664,552		$1,642,572	$1,584,692		$1,271,184	$2,082,105	
		B) BAB			F) CMA BAB			J) Δ Personal Savings Rate BAB			N) Δ Median Sales Prc for New Houses BAB	
Excess Return	0.81%	0.90%	0.79%	0.85%	0.89%	1.02%	0.99%	0.80%	1.21%	1.00%	0.82%	1.20%
CAPM alpha	0.47%***	0.18%	0.73%***	0.19%	0.45%***	0.21%	0.40%***	0.30%***	0.58%***	0.45%***	0.27%**	0.68%***
Carhart alpha	0.30%***	0.27%***	0.40%***	0.35%***	0.18%**	0.77%***	0.38%***	0.20%***	0.73%***	0.33%***	0.27%***	0.56%***
FF5 alpha	0.22%***	0.10%	0.40%***	0.22%**	0.07%	0.81%***	0.30%***	0.03%	0.80%***	0.20%**	0.14%**	0.53%***
Sharpe Ratio	0.22	0.12	0.26	0.12	0.19	0.10	0.16	0.15	0.16	0.17	0.14	0.20
Size	$1,648,315	$1,566,048		$1,882,176	$1,540,626		$1,379,864	$1,759,873		$1,522,384	$1,543,852	
		C) SMB BAB			G) Inflation BAB			K) Δ Real Personal Consumption BAB			O) Δ New One Family Houses Sold BAB	
Excess Return	0.75%	1.01%	0.62%	0.92%	0.87%	1.06%	0.86%	0.95%	0.91%	0.93%	0.88%	1.06%
CAPM alpha	0.37%***	0.32%**	0.48%***	0.38%***	0.30%**	0.56%***	0.30%**	0.41%***	0.35%**	0.37%***	0.33%***	0.51%***
Carhart alpha	0.28%***	0.32%***	0.34%***	0.34%***	0.23%***	0.60%***	0.21%***	0.35%***	0.32%**	0.29%***	0.28%***	0.52%***
FF5 alpha	0.15%***	0.22%**	0.21%*	0.19%***	0.12%***	0.52%***	0.12%*	0.20%***	0.39%***	0.20%***	0.10%	0.59%***
Sharpe Ratio	0.20	0.13	0.20	0.16	0.14	0.17	0.15	0.17	0.14	0.16	0.15	0.16
Size	$3,691,528	$633,444		$1,471,133	$1,810,149		$1,470,632	$1,635,500		$1,349,336	$1,853,359	
		D) HML BAB			H) Δ Unemployment BAB			L) Δ Average Hourly Earnings BAB			P) Default Spread BAB	
Excess Return	0.75%	1.00%	0.79%	0.99%	0.84%	1.20%	0.97%	0.86%	1.13%	1.02%	0.80%	1.31%
CAPM alpha	0.12%	0.53%***	0.05%	0.45%***	0.28%**	0.69%***	0.43%***	0.29%**	0.65%***	0.48%***	0.22%*	0.82%***
Carhart alpha	0.33%***	0.23%***	0.69%***	0.42%***	0.18%**	0.76%***	0.30%***	0.30%***	0.46%***	0.50%***	0.10%	0.99%***
FF5 alpha	0.25%***	0.07%	0.84%***	0.27%***	0.07%	0.69%***	0.14%**	0.21%***	0.37%***	0.28%***	0.06%	0.76%***
Sharpe Ratio	0.11	0.19	0.08	0.17	0.14	0.20	0.17	0.14	0.20	0.18	0.13	0.21
Size	$1,951,898	$1,574,704		$2,006,288	$1,188,775		$1,460,885	$1,620,558		$1,896,672	$1,223,148	

Table 6: Significant t-statistics from all possible univariate regressions of the levered strategies

This table shows the pricing errors' significance levels obtained from regressing each of the sixteen levered strategies on the other nineteen individually. Significance levels are calculated using heteroskedastic robust standard errors. Rows are assigned to the strategies used as dependent variables, while columns are assigned to the strategies used as regressors. For example, the first row corresponds to the significance level of the alpha's t-statistics obtained using the Market BAB strategy as the dependent variable and all other strategies as independent. The sixteen strategies are described in Table 2. The Low, High, and Low-High portfolios used to construct the strategies are weighted with the same weights across strategies. Betas and alphas used to assign the assets to the low and high portfolios are estimated with OLS. I use 5-year data to calculate betas and alphas. Portfolios are rebalanced yearly, at the end of December. I use monthly data from January 1968 – December 2015 to construct portfolios for the period January 1973 – December 2015. Individual data on stock returns comes from the CRSP database, while the data for the CAPM, Carhart, and FF5 models comes from Kenneth French's webpage. The macroeconomic variables are used with a 6-month lag to guarantee that revised series are available when the econometrician runs the regression. The macroeconomic variables are from the Federal Reserve Bank of St. Louis webpage.

	Independent Variable															
Dependent Variable	BAB	SMB BAB	HML BAB	RMW BAB	CMA BAB	BAA	Inflation BAB	Δ Unemp BAB	Δ Labor Part BAB	Δ Pers Savings BAB	Δ Real Pers Cons BAB	Δ Avg Hour Earn BAB	Δ New Housing Auth BAB	Δ New Houses Prc BAB	Δ New Houses Sold BAB	Default Premium BAB
BAB	--	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
SMB BAB		--	***	***	***	***	***	***	***	***	***	***	***	***	***	***
HML BAB	*		--	*	*	**	*	***	**	***		*	***	***	***	**
RMW BAB	**	**	***	--	**			**					*			
CMA BAB		*	**		--	**		***	**	***		*	**	***	**	**
BAA	***	***	***	***	***	--	***	**	***	***	***	**	***	**	***	**
Inflation BAB	**	**	***	***	***		--				**					
Δ Unemployment BAB	***	***	***	***	***		**	--	***	***	***		**		***	
Δ Labor Part BAB	**	**	***	**	***				--		**					
Δ Pers Savings BAB	**	**	***	**	***					--	*					
Δ Real Pers Cons BAB		**	***	***	**						--			**		
Δ Avg Hour Earn BAB	**	***	***	***	***		**		***	***	***	--	**		***	
Δ New Housing Auth BAB	**	**	***	***	***						***		--			
Δ New Houses Prc BAB	**	***	***	***	***		**		***	***	***		**	--	***	
Δ New Houses Sold BAB	**	**	***	**	***						**				--	
Default Premium BAB	***	***	***	***	***	**	***	*	***	***	***	**	***	**	***	--

* 10%, ** 5%, *** 1%

Table 7: Results from regressing levered strategies onto models containing BAB and/or BAA

The table reports the parameters estimated using models containing BAA and/or BAB as regressors: BAB, CAPM+BAB, Carhart+BAA, and BAA+BAB. The parameters reported are the estimated alphas. Significance levels are calculated using heteroskedastic robust standard errors. The additional fourteen strategies used as dependent variable are described in Table 2. The Low, High, and Low-High portfolios used to construct the strategies are weighted with the same weights across strategies. Betas and alphas used to assign the assets to the low and high portfolios of BAA and BAB are estimated with OLS. I use 5-year data to calculate betas and alphas. Portfolios are rebalanced yearly, at the end of December. I use monthly data from January 1968 – December 2015 to construct portfolios for the period January 1973 – December 2015. Individual data on stock returns comes from the CRSP database, while the data for the CAPM, Carhart, and FF5 models comes from Kenneth French's webpage. The macroeconomic variables are used with a 6-month lag to guarantee that revised series are when the econometrician runs the regression. The macroeconomic variables are from the Federal Reserve Bank of St. Louis webpage.

	BAB	CAPM+BAB	Carhart+BAA	BAA+BAB
SMB BAB	0.08%	0.01%	0.37%***	0.08%
HML BAB	0.92%*	0.45%	0.02%	-0.23%
RMW BAB	1.30%**	0.82%**	0.09%	-0.04%
CMA BAB	1.22%**	0.69%*	0.12%	-0.08%
Inflation BAB	0.76%**	0.45%*	0.13%	-0.06%
Δ Unemployment BAB	0.93%***	0.61%***	0.37%	0.14%
Δ Labor Part BAB	0.75%**	0.39%*	-0.01%	-0.12%
Δ Pers Savings BAB	1.10%**	0.71%**	0.23%	0.11%
Δ Real Pers Cons BAB	0.54%	0.20%	-0.06%	-0.28%
Δ Avg Hour Earn BAB	0.65%**	0.36%**	0.11%	-0.09%
Δ New Housing Auth BAB	0.85%**	0.51%**	0.23%	0.03%
Δ New Houses Prc BAB	0.76%**	0.46%**	0.11%	-0.03%
Δ New Houses Sold BAB	0.73%**	0.39%*	0.14%	-0.07%
Default Premium BAB	1.08%***	0.78%***	0.71%***	0.36%*

Figure 1: Levered strategies' monthly Sharpe Ratios relative to the Market factor

This figure shows the monthly Sharpe Ratio of each of the sixteen levered strategies analyzed in this paper relative to the monthly Sharpe Ratio of the CAPM's Market factor. The sixteen strategies are described in Table 2. The Low, High, and Low-High portfolios used to construct the strategies are weighted with the same weights across strategies. Betas and alphas used to assign the assets to the low and high portfolios are estimated with OLS. I use 5-year data to calculate betas and alphas. Portfolios are rebalanced yearly, at the end of December. I use monthly data from January 1968 – December 2015 to construct portfolios for the period January 1973 – December 2015. Individual data on stock returns comes from the CRSP database, while the data for the CAPM, Carhart, and FF5 models comes from Kenneth French's webpage. The macroeconomic variables are used with a 6-month lag to guarantee that revised series are available at the moment in which the econometrician runs the regression. The macroeconomic variables are from the Federal Reserve Bank of St. Louis webpage.

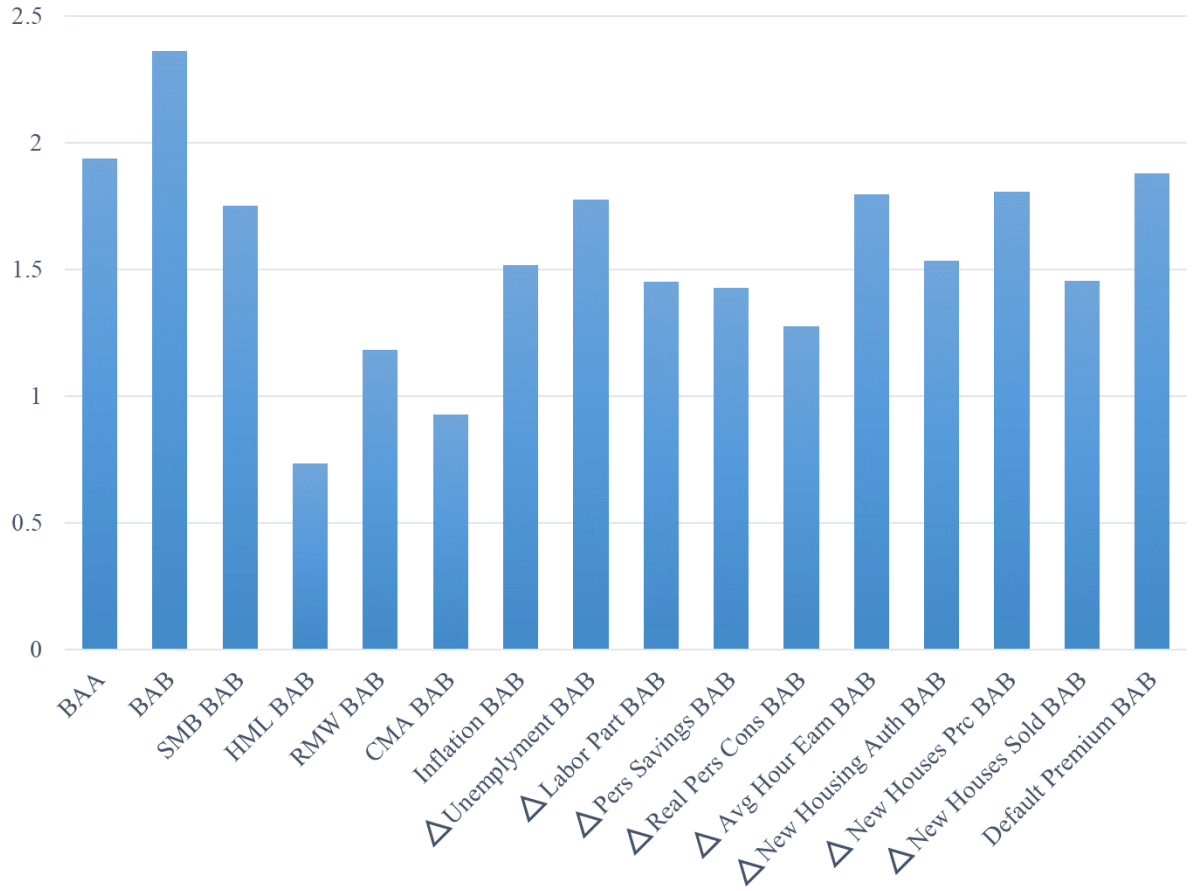
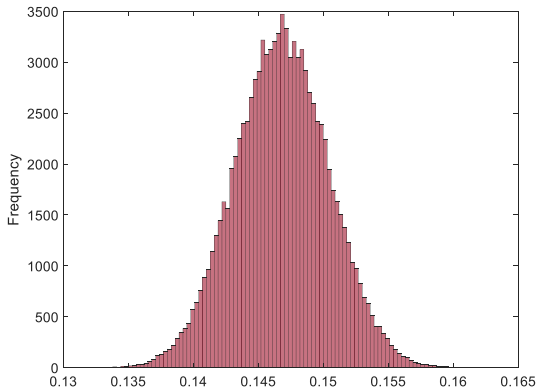


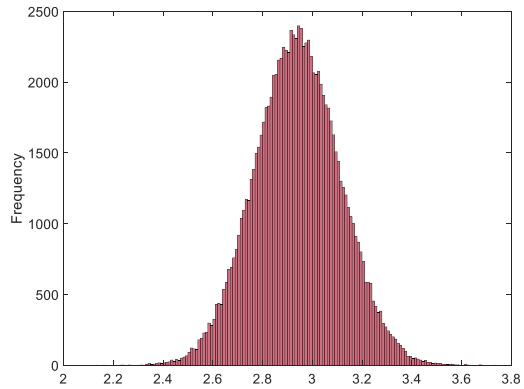
Figure 2: Distribution of 100,000 simulated Sharpe ratios and alpha t-statistics of simulated BAB strategies

This figure presents the empirical distribution of different performance metrics calculated using 100,000 simulated levered strategies. The performance metrics are the Sharpe Ratio [Panel (a)] and the heteroskedastic robust t-statistics generated by regressing each simulated strategy on the following three models: CAPM, Carhart, and FF5 [Panels (c) to (d)]. The simulated strategies are generated by randomly assigning half of the stocks to a "low" portfolio and the other half to a "high" portfolio at the end of every December. Then, I randomly assign each asset's i rank zl_i (zh_i) in the low (high) portfolio. The weight of an asset in a portfolio depends on its rank (see Section 2.2). I do this for data from December 1972 to December 2014, which means I generate monthly data on these random portfolios from January 1973 to December 2015. Individual data on stock returns comes from the CRSP database, while the data for the CAPM, Carhart, and FF5 models comes from Kenneth French's webpage.

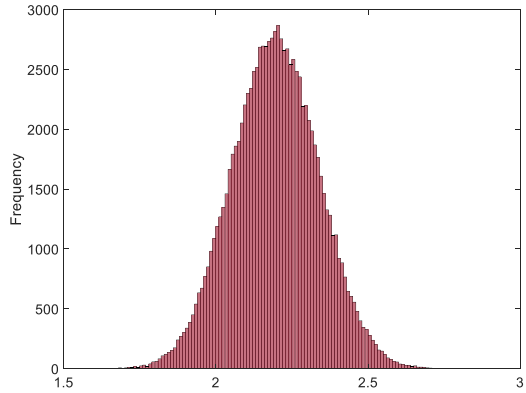
a) Sharpe ratio



c) Carhart's alpha t-statistic



b) CAPM's alpha t-statistic



d) FF5's alpha t-statistic

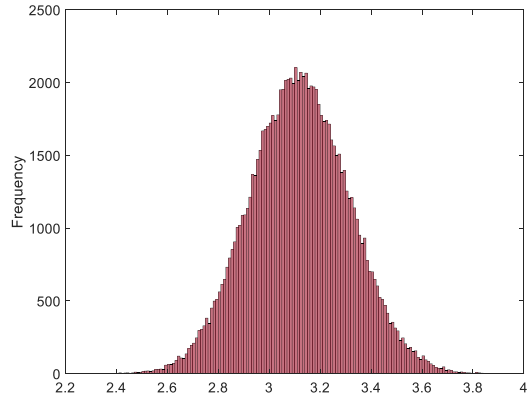
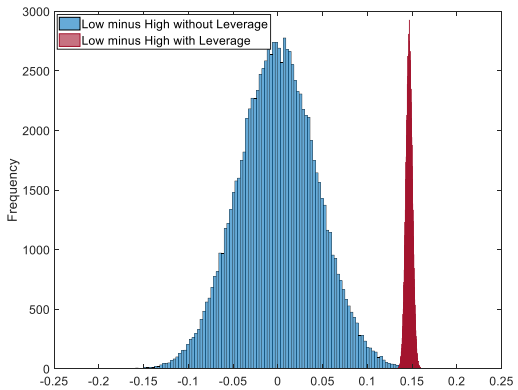


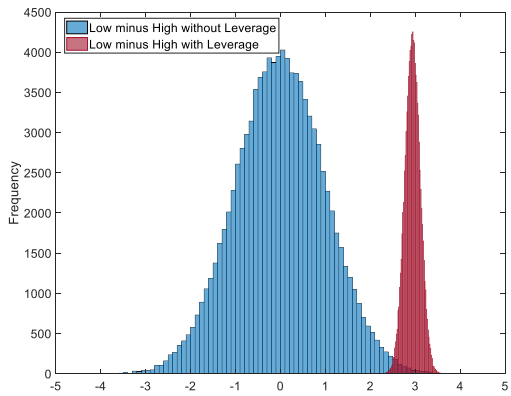
Figure 3: Comparison of Sharpe ratio and alpha t-statistics' distributions between BAB and unlevered BAB strategies

This figure presents the empirical distribution of different performance metrics calculated using 100,000 simulated strategies. Each panel plots the empirical distributions of two simulated strategies: a long-short strategy with leverage (purple) and a long-short strategy without leverage (blue). The performance metrics are the Sharpe Ratio [Panel (a)] and the heteroskedastic robust t-statistics generated by regressing each simulated strategy on the following three models: CAPM, Carhart, and FF5 [Panels (c) to (d)]. The simulated strategies are generated by randomly assigning half of the stocks to a "low" portfolio and the other half to a "high" portfolio at the end of every December. Then, I randomly assigned each asset's i rank zl_i (zh_i) in the low (high) portfolio. The weight of an asset in a portfolio depends on its rank (see Section 2.2). I do this from December 1972 to December 2014, which means I generate monthly data on these random portfolios from January 1973 to December 2015. Individual data on stock returns comes from the CRSP database, while the data for the CAPM, Carhart, and FF5 models comes from Kenneth French's webpage.

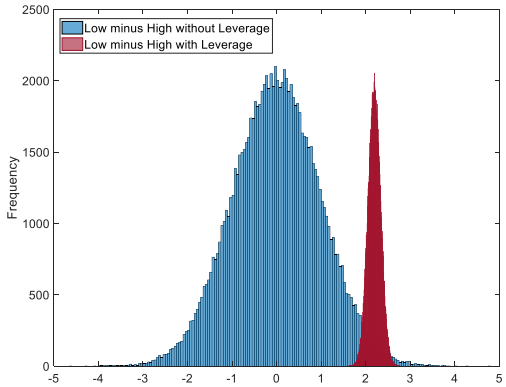
a) Sharpe ratio



c) Carhart's alpha t-statistic



b) CAPM's alpha t-statistic



d) FF5's alpha t-statistic

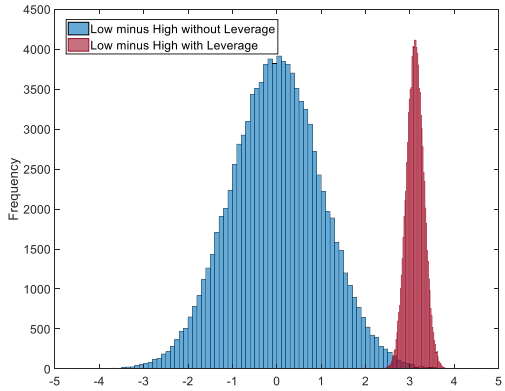
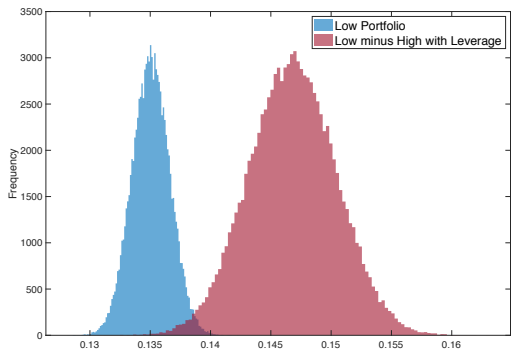


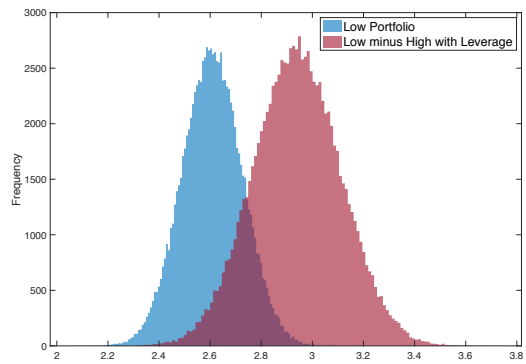
Figure 4: Comparison of Sharpe ratio and alpha t-statistics' distributions between the random BAB strategy and the excess return of the random low portfolio

This figure presents the empirical distribution of different performance metrics calculated using 100,000 simulated strategies. Each panel plots two empirical distributions: one for the long-short strategy with leverage (purple) and one for the long portfolio (low portfolio) used to construct the levered strategy (blue). The performance metrics are the Sharpe Ratio [Panel (a)] and the heteroskedastic robust t-statistics generated by regressing each simulated strategy on the following three models: CAPM, Carhart, and FF5 [Panels (c) to (d)]. The simulated strategies are generated by randomly assigning half of the stocks to a "low" portfolio and the other half to a "high" portfolio at the end of every December. Then, I randomly assign each asset's i rank zl_i (zh_i) in the low (high) portfolio. The weight of an asset in a portfolio depends on its rank (see Section 2.2). I do this for data from December 1972 to December 2014, which means I generate monthly data on these random portfolios from January 1973 to December 2015. Individual data on stock returns comes from the CRSP database, while the data for the CAPM, Carhart, and FF5 models comes from Kenneth French's webpage.

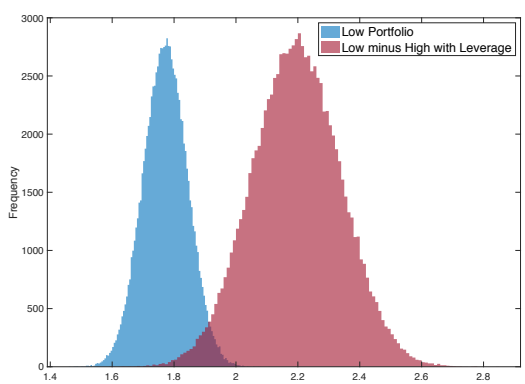
a) Sharpe ratio



c) Carhart's alpha t-statistic



b) CAPM's alpha t-statistic



d) FF5's alpha t-statistic