Decoding Systematic Relative Investing: A Pairs Approach

Christian L. Goulding
Research Affiliates, LLC, Newport Beach, CA 92660 USA

Campbell R. Harvey[*]
Duke University, Durham, NC 27708 USA
National Bureau of Economic Research, Cambridge, MA 02138 USA

Alex Pickard
Research Affiliates, LLC, Newport Beach, CA 92660 USA

Current version: December 9, 2020

We propose a novel theory that brings to light three fundamental performance drivers of zero-cost systematic investment strategies: (1) high (positive) own-asset signal-return predictability; (2) low (or negative) cross-asset signal correlation; and (3) low (or negative) cross-asset signal-return predictability. We develop these insights in the context of long–short pair strategies used as portfolio building blocks. We test our approach empirically using momentum signals for major asset classes, though our method can generalize to any signal. Our investable pairs-based portfolio harvests over double the average returns of a conventional rank-based portfolio over the last 20 years.

Keywords: cross-sectional strategies, trading strategies, long–short strategies, relative value, momentum, portfolio construction, alpha, outperformance, anomalies, market timing, asset pricing, zero-cost portfolios, pair trades

JEL Classifications: G12, G13

1

1. Introduction

A popular category of cross-sectional (CS) systematic investment strategies ranks assets each period by their signals (e.g., value, momentum, carry) and applies more weight to the assets with the highest signals and less weight to the assets with the lowest signals. These strategies have been shown to perform well for a variety of signal types and asset classes.[1] However, the top-down nature of these strategies *indirectly* implies a set of relative bets, which may or may not align well with the information structures that fuel these strategies. The signals that the strategy sorts may not only correlate with each other, but also with the subsequent returns of several other assets, and such cross-asset correlations may be stronger than own-asset correlations. Does signal sorting adequately handle these relationships? We investigate how to improve performance through a bottom-up approach, which is more focused on the relative bets it takes as informed by the structure of signal-return dependencies.
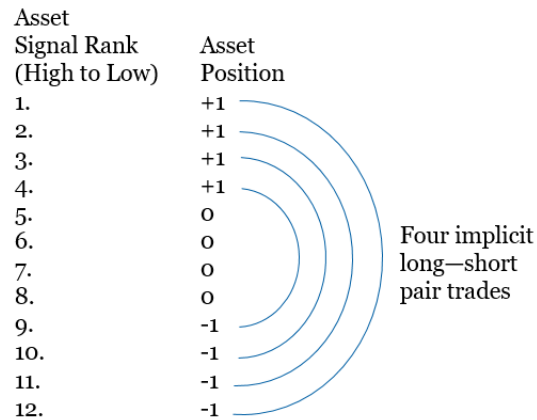
Our approach starts from a general observation: any zero-cost cross-sectional strategy on n assets can be expressed equivalently as a weighted average of long–short strategies on *pairs* of assets among these n assets.[2] For example, consider a conventional 1/3-1/3-1/3 long-neutral-short rank-based approach. Applied to 12 assets, this approach goes long the top four assets (as ranked by their signals), neutral the middle four assets, and short the bottom four assets. If we pair the 1st- and 12th-ranked assets, the 2nd- and 11th-ranked assets, and so on, then we observe four implicit long–short pair strategies embedded within this approach (see Exhibit 1). We turn this general observation around. We focus on long–short signal-based pair strategies formed from all possible asset pairs in our opportunity set, treating them as explicit building blocks rather than implicit components of a top-down approach. If we can better understand how signal-return dependencies influence the profitability of long–short pair strategies, then we can potentially use these insights to assemble portfolios with more desirable properties than standard approaches.

We postulate a general theory for long–short pair strategy performance in terms of the correlations among signals and returns of two generic assets. Our theory, which is agnostic to signal type and holding period, consists of three fundamental performance drivers: (1) own-asset signal-return predictability (correlation); (2) cross-asset signal correlation; and (3) cross-asset signal-return predictability (correlation). We argue that signal-driven long–short

---

[1]Numerous studies have documented global risk premia associated with sorting assets (such as individual securities or asset price indices) on one or more observable asset signals, including, but not limited to, market beta, size, value, momentum, profitability, investment factors, and carry (Asness et al., 2013, Fama and French, 2012 and 2015, Koijen et al., 2018, among many others).

[2]Zero cost means that the portfolio weights add to zero such that short positions finance long positions. We provide a formal statement of this observation in Section 2.

Exhibit 1: Example of Implicit Pair Trades in a Signal-Rank-Based Strategy

Asset Signal Rank (High to Low) / Asset Position:

Asset Signal Rank (High to Low)	Asset Position
1.	+1
2.	+1
3.	+1
4.	+1
5.	0
6.	0
7.	0
8.	0
9.	-1
10.	-1
11.	-1
12.	-1

Four implicit long—short pair trades

Notes: This exhibit illustrates one example of implied long–short pair trades for a signal-rank-based strategy on 12 assets. Assets are ranked by their signal levels and assigned positions according to a 1/3-1/3-1/3 long-neutral-short strategy: long the four top-ranked assets, neutral the four middle-ranked assets, and short the four bottom-ranked assets. One possible collection of equivalent long–short pair trades arises from pairing the 1st and 12th assets, the 2nd and 11th assets, the 3rd and 10th assets, and the 4th and 9th assets in long–short trades.

pair strategies perform better given high (positive) values for driver (1) and low (or negative) values for drivers (2) and (3).

We combine these three drivers into a composite function or "score" such that more desirable values for each of the three drivers corresponds to a higher composite score. Our composite function is motivated by the functional form of the expected return arising from a stylized auxiliary model. This model considers a generic asset pair whose returns are governed by linear own-asset and cross-asset signal relationships in a framework of jointly normally distributed signals and uncertainty. We use this composite function as a measure of relative profitability in a collection of long–short pair strategies.

We test the predictive merit of our three drivers and our composite score using momentum-based signals for a broad set of 13 major global asset indices, which permits 78 different possible pairings.[3] The assets include a variety of bond and stock indices as well as commodities and real estate indices. We split our sample roughly in half into an early period for estimation and detection of profitable pair strategies using our composite score, and a late period for measuring out-of-sample (OOS) performance. In the cross-section of pair strategies, each of the three drivers and the composite score exhibit a highly statistically significant predictive relationship with OOS performance, consistent with the directions in-

---

[3]Our theoretical framework is agnostic to signal type (e.g., momentum, value, carry). However, our empirical analysis focuses on momentum-based signals as described in Section 3.

3

dicated by our theory. The composite score is a stronger predictor than any of its component drivers taken individually, lending support to the functional form of the composite score as motivated by the auxiliary model.

We also find that our composite score is a stronger predictor of future pair strategy performance than is past pair strategy performance. The predictive power of past performance is not significant after controlling for the composite score, but the composite score remains significant after controlling for past performance. This evidence supports the hypothesis that the composite score bypasses some noise in pair strategy performance unrelated to the underlying drivers revealed by our theory.

We use our insights to construct a portfolio consisting of pair strategies with the highest composite scores. We estimate signal-return correlations and compute the composite score for every possible pair strategy using a rolling estimation window prior to portfolio formation.[4] In our running example, each month we construct a portfolio using 12-month momentum for each asset's current signal and a 10-year estimation window for each pair's composite score. We include the top four pair strategies (ranked by composite score). Our choice of using the top four pair strategies is motivated in part by the fact that a conventional 1/3-1/3-1/3 long-neutral-short portfolio on the 13 assets implies four long–short pairs each month (four assets long, four assets short, and the remaining five assets neutral). Because we estimate composite scores on a rolling basis, the portfolio adaptively includes or excludes asset pairs based on profitability potential as indicated by their estimated scores. During the 20-year OOS evaluation period (2000-01 to 2020-05), 20 different pairs and all 13 assets are active in the portfolio at some time.

The pairs-based portfolio provides improvements on several dimensions compared to a conventional 1/3-1/3-1/3 rank-based strategy.[5] The pairs approach harvests over 2.5 times the average return, more than doubles the Sharpe ratio, more than doubles the alpha to US equities, and increases cumulative returns during major crisis periods (2008–2009 and 2020-01 to 2020-05). The worst drawdown and worst monthly return are less severe. The pairs approach also incurs less turnover. This relative outperformance is consistent in subsamples and robust to parameter choices, such as the momentum horizon for signals, the estimation window for composite scores, and the number of top pairs included each month.[6]

---

[4]Correlations among different assets are known to change over time (Erb et al., 1994, Longin and Solnik, 1995, Boucrelle et al., 1996 and Campbell et al., 2002). Therefore, we estimate the correlations that drive our portfolio selection on a rolling basis.

[5]We use this benchmark as an illustration. There are numerous other possible candidates of top-down rank-based strategies for comparison. However, the central question is how signal-return dependencies among the assets translate to profitability, or lack thereof, for a given strategy. Therefore, this benchmark plays the role of a commonly used strategy for which the answer to this question is not directly embedded.

[6]In robustness checks in Appendix C, we analyze an array of momentum lookback horizons, estimation

4

Our method is not to be confused with the popular strategy called "pairs trading." Developed by industry researchers at Morgan Stanley in the 1980s and first analyzed by Gatev et al. (2006), pairs trading relies on detecting *temporary* price departures (widening spreads) between two historically highly correlated stocks, and then trading to take advantage of an expected narrowing of spreads as prices revert.[7] In contrast, our approach focuses on measuring systematic long-term dependence properties (correlations) of an asset pair, not only between returns, but also among signals and returns. The literature has given little attention to the function of signal-return dependence properties in portfolio choice.[8] We take advantage of these properties through sustained engagement in pair strategies that exhibit the desired properties.

Our framework is agnostic to signal type, but we focus on momentum signals for their uniformity of measurement across heterogeneous asset classes and because they tend to possess positive own-asset predictability of subsequent returns from past returns (Jegadeesh and Titman, 1993, 2001, Asness, 1994, Conrad and Kaul, 1998, Lee and Swaminathan, 2000, and Gutierrez and Kelley, 2008). Moreover, this own-asset predictability tends to be stable across heterogeneous assets and countries (Rouwenhorst, 1998, Griffin et al., 2003, Erb and Harvey, 2006, Israel and Moskowitz, 2012, and Asness et al., 2013).

Our paper relates to the vast literature on correlation and co-movement among financial assets established by Markowitz (1952) and Roy (1952) in the context of portfolio construction. Since these foundational studies, there has been much interest in cross-asset, cross-market, and cross-country correlations (Beltratti and Shiller, 1992, Engle et al., 1994, Erb et al., 1994, Longin and Solnik, 1995, Boucrelle et al., 1996, Karolyi and Stulz, 1996, Santis and Gerard, 1997, Ramchand and Susmel, 1998, Chen and He, 2001, Gulko, 2002, Cappiello et al., 2006, Mizuno et al., 2006, and Andersson et al., 2008). By applying our insights on signal-return joint dependence to a set of heterogeneous assets across different markets and countries, our application provides a new perspective on the connections between such correlations and portfolio selection.

2. Theory

Suppose we are presented with the task of choosing weights for a zero-cost portfolio on n assets. Rather viewing these n assets as our opportunity set for which we specify n portfolio

---

window lengths, number of included pairs, and weighting schemes.

[7]Further discussion and history on pairs trading can be found in Vidyamurthy (2004) and Whistler (2004).

[8]Using a different framework, Kelly et al. (2020) also point out the importance of cross-asset signal-return predictability in portfolio choice. However, our approach distinguishes the additional role of cross-signal correlation and its interaction with signal-return predictability.

weights, we can equivalently view the collection of long–short asset pairs among these n assets as our opportunity set for which we specify weights. This fact is stated formally in Proposition 1.

Proposition 1 (Zero-Cost Portfolios as Sum of Long–Short Asset Pairs). *Let*

$$r = \sum_{i=1}^{n} w_i r_i \tag{1}$$

be the return of a portfolio of $i = 1, \ldots, n$ assets ($n \geq 2$), where r_i is the return on asset i and w_i is the portfolio weight on asset i. Suppose that the portfolio has zero cost: $\sum_{i=1}^{n} w_i = 0$. Then, there exists a collection of nonnegative weights $\{\eta_{ij}\}$ on asset pairs ij for $i = 1, \ldots, n$ and $j = 1, \ldots, n$ such that the original portfolio return can be expressed equivalently as a weighted average of long–short pair returns $r_i - r_j$ among the n assets:

$$r = \sum_{i=1}^{n} \sum_{j=1}^{n} \eta_{ij}(r_i - r_j). \tag{2}$$

Proof of Proposition 1. See Appendix A. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

By Proposition 1, searching over the collection of pair weights in (2) poses no restrictions on our original task of choosing asset weights in (1). Therefore, we focus our attention on the properties of long–short pairs to derive insights for constructing a zero-cost portfolio of several assets. Note that we are not referring to asset pairs with *static* long–short positioning. Rather, we study long–short pair *strategies*, in which the long asset and short asset in a given asset pair could change through time depending on the assets' signals.

Three Drivers of Long–Short Pair Strategy Performance

We review the anatomy of a generic long–short signal-driven pair strategy. Let the assets in the pair be labeled by subscripts 1 and 2, respectively. Denote the signal on asset i at date t by $x_{i,t}$ and the return on asset i at date $t+1$ by $r_{i,t+1}$, for $i = 1, 2$. The zero-cost long–short pair strategy on assets 1 and 2 goes long one unit in the asset with the higher signal, short one unit in the asset with the lower signal, and neutral in both assets if their signals are equal. Therefore, the return at date $t+1$ on this long–short pair strategy can be expressed as follows:

$$r_{1,2,t+1} = \text{sign}(x_{1,t} - x_{2,t}) \cdot (r_{1,t+1} - r_{2,t+1}), \tag{3}$$

where the function $\text{sign}(\cdot)$ equals 1 if its argument is positive, -1 if it is negative, and 0 if it is zero. This strategy return seeks to exploit return gaps to the extent they are predicted

by signal gaps. For simplicity, we suppress time subscripts in our notation going forward.

We postulate three types of performance drivers for the long–short pair strategy of (3) based on the correlations between signals and the correlations between signals and returns. We summarize these drivers and their desirable properties in Exhibit 2.

Exhibit 2: Intuitive Drivers of Long–Short Pair Strategy Performance
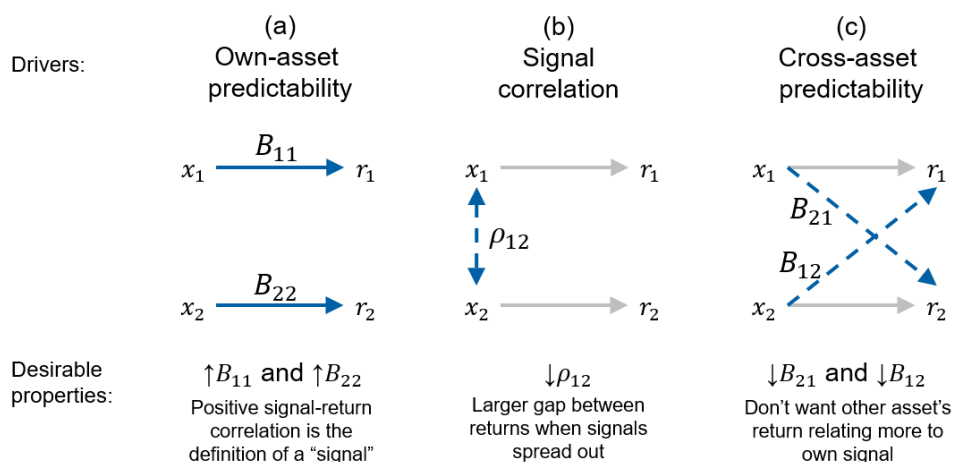
Notes: This exhibit illustrates three categories of signal-return relationships between two generic assets. Asset signal levels at date t are denoted by $x_{1,t}$ and $x_{2,t}$, respectively. Asset returns at date $t+1$ are denoted by $r_{1,t+1}$ and $r_{2,t+1}$, respectively. Our theory postulates how the signal-driven return on a long–short strategy of the two assets, $r_{1,2,t+1} = \text{sign}(x_{1,t} - x_{2,t}) \cdot (r_{1,t+1} - r_{2,t+1})$, depends on the following three categories (we suppress time subscripts in the graphic): (a) own-asset predictability, where B_{11} and B_{22} denote the correlation between each asset's signal and its subsequent return, and higher (positive) values of B_{11} and B_{22} are desirable; (b) signal correlation, where ρ_{12} denotes the cross-asset correlation between signals, and lower (negative) values of ρ_{12} are desirable; (c) cross-asset predictability, where B_{12} and B_{21} denote the correlation between each asset's signal and the other asset's subsequent return, and lower (negative) values of B_{12} and B_{21} are desirable.

Driver 1: Own-Asset Predictability

Consider the correlation between an asset's signal and its own subsequent return—a predictive relationship from t to $t+1$. We denote these correlations for assets 1 and 2 by B_{11} and B_{22}, respectively. See Exhibit 2(a). By definition of x_1 as asset 1's "signal", x_1 should positively correlate with asset 1's subsequent return, r_1. Higher values of x_1 will tend to indicate higher values of r_1, although this tendency may be noisy. Likewise, the same relationships should hold for asset 2. The higher these correlations are, the more information the signals provide and, therefore, the better our ability to use differences in the levels of these signals to separate differences in the returns of these assets.

7

Driver 2: Signal Correlation

Consider the correlation between asset signals—a contemporaneous relationship. We denote this correlation for assets 1 and 2 by ρ_{12}. See Exhibit 2(b). Are higher or lower levels of ρ_{12} desirable for the profitability of the pair strategy? Assume that Driver 1 is positive for both assets.

Case 1: $\rho_{12} > 0$. When one signal is high, the other signal tends to be high. When both signals are high, both returns tend to be high, given that signals are positive predictors of returns. When both returns are high, even if the long–short bet is in the right direction, the gap between returns will be relatively small compared with an unconditional average outcome. Likewise, when one signal is low, the other signal tends to be low, and both returns tend to be low, again exhibiting a relatively smaller return gap for the long–short bet to harvest.

Case 2: $\rho_{12} < 0$. When one signal is high, the other signal tends to be low. The higher-signal asset will tend to have a higher return while the lower-signal asset will tend to have a lower return, given that signals are positive predictors of returns. The gap between returns will be relatively large compared with an unconditional average outcome. Likewise, when one signal is low, the other signal tends to be high, and the returns tend to separate, again resulting in a relatively larger return gap for the long–short bet to harvest.

Case 3: $\rho_{12} \approx 0$. The level of one signal has relatively little or no connection to the level of the other signal. In turn, the return on the long asset tends to have little connection to the return on the short asset. Therefore, return gaps tend to be relatively noisier: not as consistently small as in Case 1 nor as consistently large as in Case 2. This represents an intermediate return potential.

In conclusion, negative (or low positive) signal correlation, ρ_{12}, corresponds to pair strategies that are more profitable.

Driver 3: Cross-Asset Predictability

Consider the correlation between an asset's signal and the subsequent return of the *other* asset—a predictive relationship from t to $t + 1$. We denote these correlations by B_{12} (the correlation between r_1 and x_2) and B_{21} (the correlation between r_2 and x_1). See Exhibit 2(c). Are higher or lower levels of B_{12} and B_{21} desirable for the profitability of the pair strategy? Assume that Driver 1 is positive for both assets. By symmetry, we focus on B_{21} and similar considerations can be applied to B_{12}.

Case 1: $B_{21} > 0$. When x_1 is high, not only will r_1 tend to be high because of own-asset predictability, but r_2 will also tend to be high because of cross-asset predictability. Therefore,

the return gap will tend to be relatively small. In fact, if this cross-asset correlation is stronger than own-asset correlation, there is an increased tendency for the return gap to not just be small, but to be in the wrong direction relative to the long–short bet indicated by the signals.

Note that correlation is not transitive, that is, if x_1 is positively correlated with both r_1 and r_2, then r_1 and r_2 are not necessarily positively correlated with each other. Accordingly, even if x_2 is low, and therefore r_2 has a tendency to be low because of own-asset predictability, cross-asset predictability coming from x_1 could predict a higher value for r_2 than predicted by x_2 without the cross-asset effect.

Case 2: $B_{21} < 0$. When x_1 is high, r_1 will tend to be high because of own-asset predictability, while r_2 will tend to be low because of negative cross-asset predictability. Therefore, the return gap will tend to be relatively large, which supports more-profitable outcomes.

Case 3: $B_{21} \approx 0$. The level of x_1 has relatively little or no connection to the level of r_2. Therefore, return gaps tend to be relatively noisier: not as consistently small as in Case 1 nor as consistently large as in Case 2.

In conclusion, negative (or low positive) cross-asset correlations, B_{12} and B_{21}, correspond to pair strategies that are more profitable.

Combining the Drivers into a Composite Score

In this section, we quantify the profitability of a pair strategy by a single number that is a composite function, or "score," of the three drivers. As a guide for the functional form of this composite score, we develop and analyze an auxiliary model. This model and its main results are stated in Proposition 2.

Proposition 2 (Auxiliary Model)**.** *Let x_1, x_2, ε_1, and ε_2 be jointly normally distributed, where x_1 and x_2 are bivariate standard normal (mean zero, variance one) with correlation $\rho_{12} \in [-1, 1]$, and where ε_1 and ε_2 are independent of x_1 and x_2, with finite variances. Let returns r_1 and r_2 be the following linear functions of these random variables:*

$$r_1 = \mu_1 + A_{11}x_1 + A_{12}x_2 + \varepsilon_1, \tag{4}$$

$$r_2 = \mu_2 + A_{21}x_1 + A_{22}x_2 + \varepsilon_2, \tag{5}$$

for some scalars $\mu_1, \mu_2, A_{11}, A_{12}, A_{21}, A_{22}$. Let the return of the signal-driven long–short pair strategy be the following:

$$r = (r_1 - r_2) \cdot sign(x_1 - x_2). \tag{6}$$

9

Then, the expected return of this pair strategy is the following:

$$\mathbf{E}[r] = [(A_{11} - A_{12}) + (A_{22} - A_{21})]\sqrt{\frac{1 - \rho_{12}}{\pi}}. \tag{7}$$

Furthermore, $\frac{\partial \mathbf{E}[r]}{\partial A_{ii}} > 0$ and $\frac{\partial \mathbf{E}[r]}{\partial A_{ij}} < 0$ if $\rho_{12} < 1$, and $\frac{\partial \mathbf{E}[r]}{\partial \rho_{12}} < 0$ if $[(A_{11}-A_{12})+(A_{22}-A_{21})] > 0$.

Proof of Proposition 2. See Appendix A. □

The linear structure in (4) and (5) allows for cross-correlation of the signals and cross-predictability of the signals with subsequent returns as well as own-asset predictability. For example, if signal x_1 is a positive predictor of own-asset return r_1, but a negative predictor of cross-asset return r_2, then parameter A_{11} would be positive and parameter A_{21} would be negative. The auxiliary model assumes the part of each asset's return that is unexplained by the signals (ε_i) is jointly normal with, and independent of, the signals. This assumption implies a bivariate normal distribution for the asset returns. It also implies a bivariate normal distribution for the r's and x's and for the difference between the r's and the difference between the x's. This simplifies the analysis of the nonlinear transformation of these random variables in (6).

There are several notable features of the results of Proposition 2. First, the expected return of the long–short pair portfolio (7) does not depend on the mean returns of the underlying assets used in the strategy ($\mathbf{E}[r_1] = \mu_1$ and $\mathbf{E}[r_2] = \mu_2$). This result is intuitive because both signals have the same marginal distribution, and therefore, either asset is equally likely to be the long or short asset. On average, the difference in mean returns between assets washes out. This property aligns with the development of our three drivers, which did not invoke sensitivity to the assets' mean returns. Second, the expected long–short portfolio return does not depend on the covariance structure of the unexplained portions of the returns (i.e., the distribution of the ε's). This result is intuitive as well for the same reasons as above. Again, this property aligns with the development of our three drivers.

Third, the expected return (7) has a larger magnitude when the underlying signals of the two assets are less positively correlated (or more negatively correlated), which aligns with our cross-correlation driver (Driver 2). Fourth, the expected return is higher when the underlying signals are better (positive) predictors of own-asset rather than cross-asset returns, which aligns with our own-asset and cross-asset predictability drivers (Drivers 1 and 3).

Parameter A_{ij} in the auxiliary model of Proposition 2 can be viewed as a marginal effect, or beta, of subsequent return i with respect to signal j, whereas B_{ij} discussed earlier captures their correlation. Because betas and correlations reflect similar joint dependence properties

10

(covariances), which differ only in scale, we retain the functional form of (7) and substitute B_{ij}'s for A_{ij}'s as a way to link this model back to our main theory. This substitution combines the key quantities of our theory $(B_{11}, B_{12}, B_{21}, B_{22}, \rho_{12})$ into a composite score (8) for measuring the relative profitability of a pair strategy.

Definition 3 (Composite Score). *Define*

$$\theta := [(B_{11} - B_{12}) + (B_{22} - B_{21})]\sqrt{\frac{1 - \rho_{12}}{\pi}}. \tag{8}$$

By inspection of (8), the marginal sensitivities of our composite score θ to each of B_{11}, B_{12}, B_{21}, B_{22}, and ρ_{12} matches the desired properties discussed earlier. Moreover, θ is a proxy for the expected return of the pair strategy. Specifically, if we view the correlations B_{11}, B_{12}, B_{21}, B_{22}, and ρ_{12} as functions of the underlying auxiliary model parameters A_{11}, A_{12}, A_{21}, A_{22}, and ρ_{12}, then the equation for θ has the same marginal sensitivities to the underlying parameters of that model as does the expected pair strategy return in (7)—under mild sufficient conditions on the model parameters.[9] This result is stated in Proposition 4.

Proposition 4 (Composite Score Marginal Sensitivities). *If the signal coefficients for asset i are of the same sign $(sign[A_{ii}] = sign[A_{ij}])$ or are sufficiently small in magnitude—specifically, if $|A_{ii}A_{ij}| < \frac{1}{2}\mathbf{Var}[\varepsilon_i]$—for $i \neq j$, then the partial derivatives of θ with respect to each of A_{11}, A_{12}, A_{21}, A_{22}, and ρ_{12} have the same signs, respectively, as the partial derivatives of $\mathbf{E}[r]$ with respect to each of A_{11}, A_{12}, A_{21}, A_{22}, and ρ_{12}.*

Proof of Proposition 4. See Appendix A. □

The implication of Proposition 4 is that any change to a model parameter that causes the expected long–short strategy return to increase also causes the value of θ to increase and vice versa, thereby providing analytical support for the functional form of our composite score.

Of course, the auxiliary model relies on several assumptions that are unlikely to hold in real markets, such as jointly normal signals and returns with stationary distributions. Empirical return distributions tend to be nonstationary and exhibit heavier tails with stronger co-movements than can be captured by the jointly normal distribution. In addition, limited liability rules out normality for marginal return distributions at the lower tail, however, the arguments in support of the intuitive drivers of our theory do not appeal to these assumptions. We use the auxiliary model as motivation for the functional form of θ.

---

[9]The conditions in Proposition 4 represent a restriction on high multicolineary among the signals and returns: cross-products between opposite-sign signal coefficients are not too large in magnitude.

Ultimately, the extent to which our three intuitive drivers, as combined in our composite score, can distinguish the profitability of different pair strategies is an empirical question. We carry out empirical assessments next.

3.   Empirical Tests of the Composite Score

In this section, we perform empirical tests of our theory using all possible 78 pair strategies formed on an opportunity set consisting of 13 broad asset-class indices.

Data

We use monthly returns for each of the following asset class indices starting from the dates listed below through 2020-05:

1. Short-Term Bonds (Barclays US Treasury 1–3yr, 1992-02),
2. US Core Bonds (Barclays US Aggregate, 1976-02),
3. Global Core Bonds (Barclays Global Agg ex-US, Hedged, 1990-02),
4. Long Treasury Bonds (Barclays US Treasury 10–20yr, 1992-02),
5. Long Credit Bonds (Barclays US Long Credit, 1973-02),
6. Emerging Markets (EM) USD Bonds (JPMorgan EMBI+, Unhedged, 1994-01),
7. High Yield (Barclays Corporate High Yield, 1983-08),
8. Commodities (Bloomberg Commodity Index, 1991-02),
9. REITs (FTSE NAREIT ALL REITS, 1972-01),
10. US Equities Small (Russell 2000, 1979-01),
11. US Equities (S&P 500, 1970-01),
12. Dev ex US Equities (MSCI EAFE, 1970-01), and
13. Emerging Markets (EM) Equities (MSCI EM (Emerging Markets), 1988-01).

Signals

Whereas our theoretical framework is agnostic to signal type (e.g., momentum, value, carry) and holding period (e.g., monthly, weekly), for empirical analysis we focus on momentum-based signals at the monthly trading frequency. Momentum signals are straightforward to estimate and apply across heterogeneous assets. For each asset i, each month we estimate its raw current momentum, s_i, as the (arithmetic) average of its trailing monthly returns over lookback horizon *mom* (in months).

For heterogeneous assets, such as the ones in our asset universe, the means and standard deviations of momentum signals can vary substantially. Using these raw momentum signals

12

to determine the long and short assets in a pair can inject direction biases. For example, stocks tend to have higher trailing return levels than bonds. Therefore, in a stock–bond pair, raw signals would indicate a long position for stocks a majority of the time whether or not that is effective. Accordingly, we employ a standardized momentum signal, x_i, for each asset i as follows:

$$x_i = \frac{s_i - \mu_i}{\sigma_i}, \tag{9}$$

where μ_i and σ_i are the inception-to-date average and standard deviation, respectively, of raw signal s_i. Note that this standardization of signals is consistent with the auxiliary model in Proposition 2, which assumes that each signal, x_i, has zero mean and unit variance.

Going forward, our analysis focuses on 12-month momentum signals. We obtain similar results in robustness checks using pair strategies governed by other lookback horizons $(1, 2, \ldots, 11 \text{ months})$.[10]

Fundamental Drivers as Performance Predictors

We split our sample roughly in half into early and late periods for the purpose of assessing the predictive power of the performance drivers of the theory.[11] In the early period, we use pre-2000 data to estimate θ and its components, as specified in (8), for each of the 78 pair strategies. Each component is a signal-signal or signal-return correlation, which we estimate using standard sample correlation statistics. The later period, 2000–forward, we set aside for testing out-of-sample (OOS) performance.

We find that each of the three drivers exhibits a highly statistically significant predictive relationship with OOS pair strategy performance in the cross section, consistent with the theory.[12] First, the cross-sectional (CS) correlation between own-asset predictability ($B_{11} + B_{22}$) and OOS average pair strategy return is $+0.27$. Second, the CS correlation between signal correlation (ρ_{12}) and OOS average pair strategy return is -0.42. Third, the CS correlation between cross-asset predictability ($B_{12} + B_{21}$) and OOS average pair strategy return is -0.28. Each of these correlations is statistically different from zero at the 1% significance level with its sign consistent with the prediction of the theory. Furthermore, the composite score, θ, exhibits a positive predictive relationship with pair strategy performance. The CS correlation of the composite score with OOS average return is $+0.44$, which is of larger magnitude than any of its component drivers taken individually. This fact lends

---

[10]See Appendix C.

[11]In our portfolio application, we generalize estimation periods to rolling windows of various lengths.

[12]Empirical in-sample relationships between the fundamental drivers and pair strategy performance are also consistent with the theory, but we emphasize the *predictive* relationships because our goal is to use this theory to form investable strategies.

support to the functional form of θ as motivated by the auxiliary model.

Composite Scores and Past Performance

The composite score is a better predictor of future pair-strategy performance than is past pair performance. In Exhibit 3, we report the estimates of three cross-sectional models. In Model 1, we regress future pair-strategy performance (post-2000) against its composite score (pre-2000). We find that the composite score is a highly significant predictor of future performance, having a t-statistic of 5.1, and $R^2 = 0.14$. In Model 2, we regress future pair strategy performance (post–2000) against past performance (pre–2000). We find that past performance is also a significant predictor of future performance, having t-statistic of 2.8. This result is not unexpected given that the composite score is based on a theoretical model of expected pair-strategy performance. In Model 3, however, in which we include both the composite score and past performance as explanatory variables, we find that the composite score remains highly statistically significant, while past performance is no longer statistically significant. We also find that the adjusted R^2 exhibits almost no increase relative to Model 1.

Exhibit 3: Cross-Sectional Regression of Post-2000 Pair Strategy Performance on Pre-2000 Composite Score θ and Past Pair Performance

	Model 1	Model 2	Model 3
Constant	0.005	0.011	0.004
	(1.28)	(3.31)	(0.95)
Composite Score θ (Pre-2000)	0.105		0.097
	(5.14)		(4.31)
Past Pair Performance (Pre-2000)		0.129	0.043
		(2.78)	(0.88)
Adjusted R^2	0.14	0.04	0.14
Number of Observations	78	78	78

Notes: This exhibit reports the estimates of three cross-sectional models across 78 possible pair strategies (for 13 assets), with t-statistics in parentheses below point estimates. Each model regresses post-2000 pair strategy performance against one or two pre-2000 pair variables: pair composite score and/or past pair performance. Pair performance is measured as the annualized average monthly return of the signal-driven long–short strategy on the pair. The pair composite score is $\theta = (B_{11} - B_{12} + B_{22} - B_{21})\sqrt{\frac{1-\rho_{12}}{\pi}}$, where B_{ij} is the sample correlation between the monthly signal on asset j and the subsequent monthly return on asset i, for $i = 1, 2$ and $j = 1, 2$, and ρ_{12} is the sample correlation between the signals. Signals are standardized 12-month momentum.

This evidence suggests that the composite score can extract meaningful information about pair strategy performance above and beyond past performance itself. Moreover, it suggests that the composite score bypasses some noise in pair strategy performance unrelated to the underlying performance drivers indicated by the theory.

4. Portfolio Application

In this section, we apply the theory of Section 2, using momentum signals on the universe of 13 assets detailed in Section 3, to assemble a bottom-up pairs-based CS portfolio and compare it with a standard top-down rank-based CS portfolio.

Pairs-Based Portfolios

We consider a family of portfolios assembled from the 78 possible pair strategies available on our 13 assets. We denote these portfolios by

$$\text{Pairs}(mom =, win =, n =, p =),$$

where mom is the momentum lookback horizon (in months) as described in Section 3 and the other parameters are defined as follows. For each asset pair ij, each month we estimate its composite score, θ_{ij}, using a rolling window of trailing signals and returns of length win months. Each estimate is based on combining the standard sample estimates of the own-asset and cross-asset correlations as they appear in (8).

Each month we rank pair strategies by their estimated θ's and select the top n pairs for inclusion in the portfolio. Note that the asset signals, x_i and x_j, in the pair ij determine which asset is long or short in that pair for the month. Therefore, even if a particular pair is in the top n pairs in consecutive months, the bets taken on the assets within that pair are not fixed, but vary with the relative levels of their signals.

Because we expect higher levels of θ to predict higher expected returns, we allow for more weight to be given to the higher-ranked pairs within the top n via a tilting parameter p. Specifically, the relative weight on each pair strategy ij each month is proportional to $(\theta_{ij})^p$, which is its estimated composite score raised to a power p, with $p \geq 0$. For example, if $p = 0$, then each of the top n pairs gets equal weight, whereas for $p > 0$, the highest ranked pair gets more relative weight than the others.[13] See Exhibit 4.

---

[13]We are fully aware that the choice of mom, win, n, and p invite a multiple testing problem detailed by Harvey et al. (2016) and Harvey (2017). As such, we are careful to provide robustness analysis to the choice of these parameters.

Exhibit 4: Example of Relative Pair Strategy Weights for Different Values of p
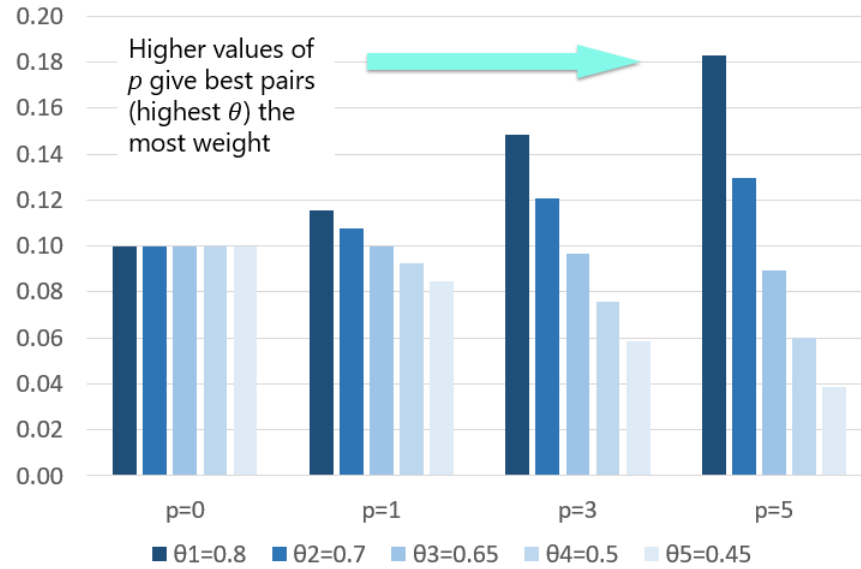
Notes: This figure illustrates the relative weights on the top pair strategies for $n = 5$ at various values of the tilting parameter p: $p = 0, 1, 3, 5$. The figure uses 0.8, 0.7, 0.65, 0.5, and 0.45 as hypothetical composite scores for the top pairs. The weights are scaled to sum to 0.5.

Benchmark Portfolios

In this section, we describe benchmark CS momentum portfolios, which we use for comparison to our pairs-based portfolios. We denote these benchmark portfolios by

$$\text{CSMOM}(mom =, LNS\% =),$$

where mom is the lookback horizon of the raw momentum signal and $LNS\%$ is the long-neutral-short split of a top-down signal-ranking algorithm. For example, $\text{CSMOM}(mom = 12, LNS\% = 1/3)$ denotes the standard top-down rank-based CS momentum portfolio, which ranks each asset each month by its raw 12-month trailing momentum signal, s_i, and goes long the top $\approx 1/3$ of assets (4/13), short the bottom $\approx 1/3$ (4/13), and neutral the middle $\approx 1/3$ (5/13).

Running Example

In remaining sections, we focus our empirical analyses on a running example, $\text{Pairs}(mom = 12, win = 120, n = 4, p = 3)$, over the evaluation period 2000-01 to 2020-05, which corre-

sponds to the post-2000 subsample of Section 3. We select this evaluation period to allow our rolling θ estimates an initial warm-up period prior to portfolio assessment. Similarly, we use the following benchmark CS portfolio: CSMOM($mom = 12, LNS\% = 1/3$). Our choice of $n = 4$ is motivated in part by the fact that the 1/3 CS momentum benchmark on 13 assets implies four long–short pairs each month (four assets long, four assets short, and the remaining five assets neutral). The parameters of our running example do not maximize backtest performance. Rather, we use this example to illustrate the effectiveness of the theory in guiding the construction of an investable pairs-based portfolio. We analyze the impact of other portfolio parameter values in robustness tests (see Appendix C).

Note the adaptive nature of pairs-based portfolios. Because we use a rolling estimation window for θ, the number of active pairs (i.e., pairs whose strategies get included in some month during our evaluation period) is likely to be higher than n. Exhibit 5 reports the number of active pairs and number of active assets (i.e., assets that are part of an active pair) as a function of n over the evaluation period. Even if we were to include only the single top-ranked pair strategy each month, there would be 8 active pair strategies. For the case of $n = 4$ in our running example, there are 20 active pair strategies and all 13 assets are active at some time. Therefore, our pairs-based portfolio has exposure to an array of assets and pair strategies over time despite setting the number of included pair strategies to the seemingly small value of $n = 4$ each month.

Historical Simulation

Consider the performance of our running example portfolio in historical simulations over the OOS evaluation period from 2000-01 to 2020-05. We scale the pairs-based and benchmark portfolios to 100% gross leverage each month (i.e., 0.5 units short in total and 0.5 units long in total).[14] Exhibit 6 shows that our pairs-based portfolio outperforms the benchmark along many dimensions. First, the average annualized monthly return of Pairs is a highly statistically significant 5.3% (3.0 t-stat.), which compares favorably to the relatively insignificant 2.0% (1.3 t-stat.) of CSMOM. The annualized Sharpe ratio of Pairs is 0.66, more than double that of CSMOM, and the alpha to US Equities is also more than doubled. Second, the worst drawdown and worst month of Pairs are less severe relative to CSMOM, and cumulative returns during crisis periods are also improved. Third, Pairs achieves this performance with annualized one-way turnover of 120%, compared to the 198% of CSMOM. Therefore, losses due to transaction costs are likely to be lower for Pairs. Note that turnover for Pairs

---

[14]Gross leverage is the sum of the absolute values of asset positions. Because the portfolio is zero cost, the net leverage, sum of asset positions, is zero.

Exhibit 5: Adaptive Nature of Pairs-Based Portfolios

Number n Top Pairs	Number of Active Pairs	Number of Active Assets
1	8	8
2	11	10
3	16	12
4	20	13
5	23	13
6	26	13
7	31	13
8	35	13
9	36	13
10	40	13
15	53	13
20	67	13

Notes: This exhibit reports the number of active pairs (out of 78) and the number of active assets (out of 13) for various values of included pair strategies, n, during the evaluation period 2000-01 to 2020-05 for the Pairs($mom = 12, win = 120$). An active pair is a pair strategy with composite score θ in the top n values for some month during the evaluation period. An active asset is an asset in an active pair during the evaluation period.

tends to be more concentrated in assets with higher liquidity, such as Short-Term Bonds, and therefore average transaction costs per unit of turnover are also likely to be lower than that of CSMOM.

In Appendix B, we analyze the composition of the pairs-based portfolio in detail, including common active pairs and the dynamics of asset weights. In Appendix C, we document the robustness of the pairs-based portfolio over subsamples and with respect to its four parameters. We find that Pairs exhibits consistent performance in subsamples, across momentum lookback horizons, and across values for n and p. We also find that for θ estimation windows centered around 120 months ($\pm$ 30 months), performance is robust and appears to balance the tradeoff between the noise that comes with shorter estimation windows and the staleness that comes with longer windows.

5. Conclusion

Signal-based portfolios of many assets are complex objects with numerous moving parts and considerations. We break down this complexity into simpler components of long–short *pair* strategies. We develop a new theory, which focuses on the joint dependence among signals and returns of a generic pair of assets. Our theory links better performance for long–short

Exhibit 6: Historical Simulation of CSMOM and Pairs Portfolios

	CSMOM	Pairs
Average Return (anlzd., %)	2.0	5.3
t-statistic	1.27	2.96
Volatility (anlzd., %)	7.3	8.1
Sharpe Ratio (anlzd.)	0.28	0.66
Skewness	0.32	0.26
Excess Kurtosis	2.79	2.00
Worst Drawdown (%)	−27.2	−22.3
Peak	2009-01	2009-01
Trough	2012-02	2010-05
Worst Month (%)	−7.7	−6.7
Month	2009-04	2001-01
Beta to US Equities	−0.16	−0.10
Alpha to US Equities (anlzd., %)	2.8	5.8
t-statistic	1.85	3.26
Cumulative Return 2008–2009 (%)	9.7	11.9
Cumulative Return 2020-01 to 2020-05 (%)	1.8	3.7
Net Leverage (%)	0	0
Gross Leverage (%)	100	100
Turnover (1-way, anlzd., %)	198	120

Notes: This exhibit reports various portfolio statistics over the evaluation period 2000-01 to 2020-05 based on monthly trading of two strategies: CSMOM($mom = 12, LNS = 1/3$) and Pairs($mom = 12, win = 120, n = 4, p = 3$).

pair strategies to three desirable properties of joint dependence: (1) high (positive) own-asset signal-return predictability; (2) low (or negative) cross-asset signal correlation; and (3) low (or negative) cross-asset signal-return predictability. We combine these properties into a single composite score to distinguish profitable pair strategies, from which we can reconstruct multi-asset signal-based portfolios.

We apply our theory empirically using momentum-based signals for a collection of broad asset-class indices, such as stocks, bonds, commodities, and real estate. First, we find that high composite scores for pair strategies are predictive of pair strategy performance out of sample and subsume the explanatory value of past pair performance for subsequent returns. Second, we assemble an investable pairs-based portfolio using the pair strategies with the top estimated composite scores each month. This portfolio harvests over double the average returns of a conventional rank-based portfolio in our out-of-sample evaluation period over the last 20 years.

Appendix A. Proofs

Proof of Proposition 1. If is sufficient to define a collection of nonnegative weights $\{\eta_{ij}\}$ such that $\sum_{j=1}^{n} \eta_{kj} - \sum_{i=1}^{n} \eta_{ik} = w_k$ for $k = 1, \ldots, n$. We define such a collection as follows. Define $\widehat{\eta}_{12} = w_1$. For $i = 2, \ldots, n-1$: let $\widehat{\eta}_{i,i+1} = \sum_{k=1}^{i} w_k$. For all other i and j, $\widehat{\eta}_{ij} = 0$. For all $i = 1, \ldots, n$ and $j = 1, \ldots, n$, if $\widehat{\eta}_{ij} \geq 0$, then $\eta_{ij} = \widehat{\eta}_{ij}$ and $\eta_{ji} = 0$; otherwise, $\eta_{ji} = -\widehat{\eta}_{ij}$ and $\eta_{ij} = 0$. Note that $\eta_{ij} = \widehat{\eta}_{ij} 1_{\{\widehat{\eta}_{ij} \geq 0\}} - \widehat{\eta}_{ji} 1_{\{\widehat{\eta}_{ij} < 0\}}$ for all i and j.

We now show for this collection that $\sum_{j=1}^{n} \eta_{kj} - \sum_{i=1}^{n} \eta_{ik} = w_k$ for $k = 1, \ldots, n$.

Case $k = 1$. $\sum_{j=1}^{n} \eta_{1j} - \sum_{i=1}^{n} \eta_{i1} = \eta_{12} - \eta_{21} = \widehat{\eta}_{12} 1_{\{\widehat{\eta}_{12} \geq 0\}} - \widehat{\eta}_{21} 1_{\{\widehat{\eta}_{12} < 0\}} - (\widehat{\eta}_{21} 1_{\{\widehat{\eta}_{21} \geq 0\}} - \widehat{\eta}_{12} 1_{\{\widehat{\eta}_{21} < 0\}}) = w_1 1_{\{\widehat{\eta}_{12} \geq 0\}} - w_1 1_{\{\widehat{\eta}_{21} \geq 0\}} = w_1$.

Case $k = 2, \ldots, n-1$. $\sum_{j=1}^{n} \eta_{kj} - \sum_{i=1}^{n} \eta_{ik} = \eta_{k,k-1} + \eta_{k,k+1} - \eta_{k-1,k} - \eta_{k+1,k} = (\eta_{k,k-1} - \eta_{k-1,k}) + (\eta_{k,k+1} - \eta_{k+1,k}) = (\eta_{k,k-1} - \eta_{k-1,k}) 1_{\{\widehat{\eta}_{k-1,k} \geq 0\}} + (\eta_{k,k-1} - \eta_{k-1,k}) 1_{\{\widehat{\eta}_{k-1,k} < 0\}} + (\eta_{k,k+1} - \eta_{k+1,k}) 1_{\{\widehat{\eta}_{k,k+1} \geq 0\}} + (\eta_{k,k+1} - \eta_{k+1,k}) 1_{\{\widehat{\eta}_{k,k+1} < 0\}} = -\widehat{\eta}_{k-1,k} 1_{\{\widehat{\eta}_{k-1,k} \geq 0\}} - \widehat{\eta}_{k-1,k} 1_{\{\widehat{\eta}_{k-1,k} < 0\}} + \widehat{\eta}_{k,k+1} 1_{\{\widehat{\eta}_{k,k+1} \geq 0\}} + \widehat{\eta}_{k,k+1} 1_{\{\widehat{\eta}_{k,k+1} < 0\}} = -\widehat{\eta}_{k-1,k} + \widehat{\eta}_{k,k+1} = -\sum_{i=1}^{k-1} w_i + \sum_{i=1}^{k} w_i = w_k$.

Case: $k = n$. $\sum_{j=1}^{n} \eta_{nj} - \sum_{i=1}^{n} \eta_{in} = \eta_{n,n-1} - \eta_{n-1,n} = \eta_{n,n-1} 1_{\{\widehat{\eta}_{n-1,n} \geq 0\}} + \eta_{n,n-1} 1_{\{\widehat{\eta}_{n-1,n} < 0\}} - \eta_{n-1,n} 1_{\{\widehat{\eta}_{n-1,n} \geq 0\}} - \eta_{n-1,n} 1_{\{\widehat{\eta}_{n-1,n} < 0\}} = -\widehat{\eta}_{n-1,n} 1_{\{\widehat{\eta}_{n-1,n} < 0\}} - \widehat{\eta}_{n-1,n} 1_{\{\widehat{\eta}_{n-1,n} \geq 0\}} = -\widehat{\eta}_{n-1,n} = -\sum_{i=1}^{n-1} w_i = -(\sum_{i=1}^{n} w_i - w_n) = -(0 - w_n) = w_n$.

$\square$

We present auxiliary lemmas, which we use to prove some propositions of the main text.

Lemma 5. *Let Z_1 and Z_2 be bivariate standard normal random variables (mean zero, variance one) with covariance τ. Then,*

$$\mathbf{E}[Z_1 | Z_2 < z] = -\tau \frac{\phi(z)}{\Phi(z)}, \tag{A1}$$

$$\mathbf{E}[Z_1 | Z_2 \geq z] = \tau \frac{\phi(z)}{1 - \Phi(z)}, \tag{A2}$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2}$ and $\Phi(z) = \int_{-\infty}^{z} \phi(\zeta) d\zeta$ are the probability density and cumulative distribution functions, respectively, of a standard normal random variable.

Proof of Lemma 5. These standard results follow from the fact that the conditional distribution of Z_1 given Z_2 is normal and from expected value calculations for truncated normal random variables. See, for example, Maddala (1983), p. 367. $\square$

Lemma 6. *Let x_1 and x_2 be bivariate standard normal random variables (mean zero, variance one) with covariance ρ_{12}. Then,*

$$\mathbf{E}[sign(x_1 - x_2)] = 0, \tag{A3}$$

$$\mathbf{E}[x_1 \cdot sign(x_1 - x_2)] = \sqrt{\frac{1 - \rho_{12}}{\pi}}, \tag{A4}$$

$$\mathbf{E}[x_2 \cdot sign(x_1 - x_2)] = -\sqrt{\frac{1 - \rho_{12}}{\pi}}. \tag{A5}$$

Proof of Lemma 6. First,

$$
\begin{aligned}
\mathbf{E}[\text{sign}(x_1 - x_2)] &= 1 \cdot P[x_1 > x_2] + 0 \cdot P[x_1 = x_2] + (-1) \cdot P[x_1 < x_2] \\
&= P[x_1 > x_2] - (1 - P[x_1 \geq x_2]) \\
&= 0,
\end{aligned}
$$

where the first equality follows from the law of total expectation, and the second and third equalities follow from the complement rule for probability. This argument proves (A3).

Second, consider two cases for $\rho_{12} \in [-1, 1]$. Case 1: $\rho_{12} = 1$. If x_1 and x_2 are perfectly positively correlated, then $\text{sign}(x_1 - x_2) = 0$ with probability 1 because x_1 and x_2 are identically distributed. Therefore, $\mathbf{E}[x_1 \cdot \text{sign}(x_1 - x_2)] = 0 = \sqrt{\frac{1 - \rho_{12}}{\pi}}$.

Case 2: $\rho_{12} \in [-1, 1)$. Let $Z_1 = x_1$ and $Z_2 = \frac{x_1 - x_2}{\sqrt{2(1 - \rho_{12})}}$. Z_2 is well defined because $\sqrt{2(1 - \rho_{12})} > 0$ for $\rho_{12} < 1$. Because x_1 and x_2 are jointly standard normally distributed, Z_2 is normally distributed with mean $\mathbf{E}[Z_2] = \frac{\mathbf{E}[x_1] - \mathbf{E}[x_2]}{\sqrt{2(1 - \rho_{12})}} = 0$ and $\mathbf{Var}[Z_2] = \frac{\mathbf{Var}[x_1] - 2\mathbf{Cov}[x_1, x_2] + \mathbf{Var}[x_1]}{2(1 - \rho_{12})} = 1$. Moreover, Z_1 and Z_2 are jointly standard normally distributed with $\mathbf{Cov}(Z_1, Z_2) = \sqrt{\frac{1 - \rho_{12}}{2}}$. Therefore,

$$
\begin{aligned}
\mathbf{E}[x_1 \cdot \text{sign}(x_1 - x_2)] &= \mathbf{E}[Z_1 \cdot \text{sign}(Z_2)] \\
&= \mathbf{P}[Z_2 > 0]\mathbf{E}[Z_1 | Z_2 > 0] - \mathbf{P}[Z_2 \leq 0]\mathbf{E}[Z_1 | Z_2 \leq 0] \\
&= \frac{1}{2}\left(\sqrt{\frac{1 - \rho_{12}}{2}}\frac{\phi(0)}{1 - \Phi(0)}\right) - \frac{1}{2}\left(-\sqrt{\frac{1 - \rho_{12}}{2}}\frac{\phi(0)}{\Phi(0)}\right) \\
&= \sqrt{\frac{1 - \rho_{12}}{\pi}},
\end{aligned}
$$

where the first equality follows from the fact that Z_2 is a positive scaling of $x_1 - x_2$; the second equality follows from the law of total expectation; the third equality follows from Lemma 5 evaluated at $z = 0$ and $\mathbf{P}[Z_2 > 0] = \mathbf{P}[Z_2 \leq 0] = \frac{1}{2}$; and the final equality follows

from substitution of $\phi(0) = \frac{1}{\sqrt{2\pi}}$ and $\Phi(0) = 1 - \Phi(0) = \frac{1}{2}$ plus simplification. This argument proves (A4).

Finally, $\mathbf{E}[x_2 \cdot \text{sign}(x_1 - x_2)] = -\mathbf{E}[x_2 \cdot \text{sign}(x_2 - x_1)] = -\mathbf{E}[x_1 \cdot \text{sign}(x_1 - x_2)]$, where the final equality follows by symmetry after interchanging subscripts. This argument proves (A5). $\qquad\square$

Proof of Proposition 2. First note that (4), (5), and (6) imply

$$r = [(\mu_1 - \mu_2) + (A_{11} - A_{21})x_1 + (A_{12} - A_{22})x_2 + (\varepsilon_1 - \varepsilon_2)] \cdot \text{sign}(x_1 - x_2).$$

Next, apply the expectation operator and Lemma 6:

$$
\begin{aligned}
\mathbf{E}[r] =\ & (\mu_1 - \mu_2) \cdot \mathbf{E}[\text{sign}(x_1 - x_2)] + (A_{11} - A_{21}) \cdot \mathbf{E}[x_1 \cdot \text{sign}(x_1 - x_2)] \\
& + (A_{12} - A_{22}) \cdot \mathbf{E}[x_2 \cdot \text{sign}(x_1 - x_2)] + \mathbf{E}[(\varepsilon_1 - \varepsilon_2) \cdot \text{sign}(x_1 - x_2)] \\
=\ & [(\mu_1 - \mu_2) + \mathbf{E}[(\varepsilon_1 - \varepsilon_2)]] \cdot \mathbf{E}[\text{sign}(x_1 - x_2)] \\
& + (A_{11} - A_{21}) \cdot \mathbf{E}[x_1 \cdot \text{sign}(x_1 - x_2)] + (A_{12} - A_{22}) \cdot \mathbf{E}[x_2 \cdot \text{sign}(x_1 - x_2)] \\
=\ & [(\mu_1 - \mu_2) + \mathbf{E}[(\varepsilon_1 - \varepsilon_2)]] \cdot 0 + (A_{11} - A_{21}) \cdot \sqrt{\frac{1 - \rho_{12}}{\pi}} + (A_{12} - A_{22}) \cdot \left(-\sqrt{\frac{1 - \rho_{12}}{\pi}} \right) \\
=\ & [(A_{11} - A_{12}) + (A_{22} - A_{21})] \sqrt{\frac{1 - \rho_{12}}{\pi}},
\end{aligned}
$$

where the first equality follows by linearity of the expectation operator; the second equality follows from the fact that $\mathbf{E}[(\varepsilon_1 - \varepsilon_2) \cdot \text{sign}(x_1 - x_2)] = \mathbf{E}[\varepsilon_1 - \varepsilon_2] \cdot \mathbf{E}[\text{sign}(x_1 - x_2)]$ by independence of ε_1 and ε_2 with x_1 and x_2; the third equality follows from application of Lemma 6; and the last equality follows by simplification of terms.

Comparative statics enumerated in the Proposition follow by inspection and complete the proof. $\qquad\square$

Proof of Proposition 4. First, we derive an expression for θ in terms of model parameters.

$$
\begin{aligned}
\theta =\ & [(B_{11} - B_{12}) + (B_{22} - B_{21})] \sqrt{\frac{1 - \rho_{12}}{\pi}} \\
=\ & [(\mathbf{Corr}[r_1, x_1] - \mathbf{Corr}[r_1, x_2]) + (\mathbf{Corr}[r_2, x_2] - \mathbf{Corr}[r_2, x_1])] \sqrt{\frac{1 - \rho_{12}}{\pi}} \\
=\ & \left[\frac{(A_{11} - A_{12})}{\sqrt{A_{11}^2 + 2\rho_{12}A_{11}A_{12} + A_{12}^2 + \mathbf{Var}[\varepsilon_1]}} + \frac{(A_{22} - A_{21})}{\sqrt{A_{22}^2 + 2\rho_{12}A_{22}A_{21} + A_{21}^2 + \mathbf{Var}[\varepsilon_2]}} \right] \frac{(1 - \rho_{12})^{\frac{3}{2}}}{\sqrt{\pi}},
\end{aligned}
$$
(A6)

where the first and second equalities follow by definition of θ and B_{ij} for $i = 1, 2$ and $j = 1, 2$; and the third equality follows from

$$\mathbf{Corr}[Y_i, X_i] = \frac{\mathbf{Cov}[A_i + A_{ii}X_i + A_{ij}X_j + \varepsilon_i, X_i]}{\mathbf{SD}[Y_i]\mathbf{SD}[X_i]} = \frac{A_{ii} + \rho_{12}A_{ij}}{\mathbf{SD}[Y_i]\mathbf{SD}[X_i]},$$

$$\mathbf{Corr}[Y_i, X_j] = \frac{\mathbf{Cov}[A_i + A_{ii}X_i + A_{ij}X_j + \varepsilon_i, X_j]}{\mathbf{SD}[Y_i]\mathbf{SD}[X_j]} = \frac{\rho_{12}A_{ii} + A_{ij}}{\mathbf{SD}[Y_i]\mathbf{SD}[X_j]},$$

for $i = 1, 2$ and $j = 1, 2$, where $\mathbf{SD}[X_i] = 1$, and $\mathbf{SD}[Y_i] = \sqrt{A_{ii}^2 + 2\rho_{12}A_{ii}A_{ij} + A_{ij}^2 + \mathbf{Var}[\varepsilon_i]}$, for $i = 1, 2$.

Next, we examine the partial derivatives of θ in (A6) with respect to model parameters A_{ij} for $(i, j) = (1, 2)$ and $(2, 1)$:

$$\frac{\partial \theta}{\partial A_{ii}} = \frac{A_{ij}(A_{ii} + A_{ij})(1 + \rho_{12}) + \mathbf{Var}[\varepsilon_i]}{(A_{ii}^2 + 2\rho_{12}A_{ii}A_{ij} + A_{ij}^2 + \mathbf{Var}[\varepsilon_i])} \frac{(1 - \rho_{12})^{\frac{3}{2}}}{\sqrt{\pi}},$$

$$\frac{\partial \theta}{\partial A_{ij}} = -\frac{A_{ii}(A_{ii} + A_{ij})(1 + \rho_{12}) + \mathbf{Var}[\varepsilon_i]}{(A_{ii}^2 + 2\rho_{12}A_{ii}A_{ij} + A_{ij}^2 + \mathbf{Var}[\varepsilon_i])} \frac{(1 - \rho_{12})^{\frac{3}{2}}}{\sqrt{\pi}}.$$

If $\rho_{12} = 1$, then $\frac{\partial \theta}{\partial A_{ii}} = \frac{\partial \theta}{\partial A_{ij}} = 0$ and $\frac{\partial \mathbf{E}[r]}{\partial A_{ii}} = \frac{\partial \mathbf{E}[r]}{\partial A_{ij}} = 0$; otherwise,

$$\text{sign}\left[\frac{\partial \theta}{\partial A_{ii}}\right] = \text{sign}[A_{ij}(A_{ii} + A_{ij})(1 + \rho_{12}) + \mathbf{Var}[\varepsilon_i]],$$

$$\text{sign}\left[\frac{\partial \theta}{\partial A_{ij}}\right] = -\text{sign}[A_{ii}(A_{ii} + A_{ij})(1 + \rho_{12}) + \mathbf{Var}[\varepsilon_i]].$$

The terms inside the sign functions are identical except in the second subscript of the parameters. Note that $(1 + \rho_{12})A_{ii}A_{ij} \geq -2|A_{ii}A_{ij}|$. Therefore, $A_{ij}(A_{ii} + A_{ij})(1 + \rho_{12}) + \mathbf{Var}[\varepsilon_i] \geq (1 + \rho_{12})A_{ij}^2 + (\mathbf{Var}[\varepsilon_i] - 2|A_{ii}A_{ij}|) > 0$, where the final inequality follows by the assumption that $\mathbf{Var}[\varepsilon_i] > 2|A_{ii}A_{ij}|$. Therefore, $\text{sign}\left[\frac{\partial \theta}{\partial A_{ii}}\right] = \text{sign}\left[\frac{\partial \mathbf{E}[r]}{\partial A_{ii}}\right] > 0$ and $\text{sign}\left[\frac{\partial \theta}{\partial A_{ij}}\right] = \text{sign}\left[\frac{\partial \mathbf{E}[r]}{\partial A_{ij}}\right] < 0$.

Finally, we examine the partial derivative of θ in (A6) with respect to model parameter ρ_{12}:

$$\frac{\partial \theta}{\partial \rho_{12}} = -\frac{1}{2}\left[(A_{11} - A_{12})K_1 + (A_{22} - A_{21})K_2\right]\sqrt{\frac{1 - \rho_{12}}{\pi}}, \tag{A7}$$

where

$$K_i = \frac{L_i(\rho_{12})}{\mathbf{SD}[r_i]^3} + \frac{\mathbf{Var}[\varepsilon_i]}{\mathbf{SD}[r_i]^3} \tag{A8}$$

23

and

$$L_i(\rho_{12}) := 3A_{i1}^2 + 2A_{i1}A_{i2}(1 + 2\rho_{12}) + 3A_{i2}^2 \tag{A9}$$

We now show that $K_i > 0$ for $i = 1, 2$. First, we show that $L_i(\rho_{12}) \geq 0$ for $\rho_{12} \in [-1, 1]$. Because $L_i(\rho_{12})$ is linear in ρ_{12}, it is minimized on $[-1, 1]$ at $\rho_{12} = 1$ or $\rho_{12} = -1$. *Case 1:* $L_i(1) \leq L_i(-1)$. This case implies $L_i(\rho_{12}) \geq L_i(1) = 3(A_{i1} + A_{i2})^2 \geq 0$. *Case 2:* $L_i(-1) \leq L_i(1)$. This case implies $A_{i1}A_{i2} \geq 0$ and, therefore, $L_i(\rho_{12}) \geq L_i(-1) = 3A_{i1}^2 - 2A_{i1}A_{i2} + 3A_{i2}^2 \geq 3A_{i1}^2 - 6A_{i1}A_{i2} + 3A_{i2}^2 = 3(A_{i1} - A_{i2})^2 \geq 0$. Both cases imply $L_i(\rho_{12}) \geq 0$. Second, because $\mathbf{Var}[\varepsilon_i] > 0$ and $\mathbf{SD}[r_i] > 0$, we have $K_i > 0$. Therefore, if $A_{11} - A_{12} > 0$ and $A_{22} - A_{21} > 0$, then $\mathrm{sign}\left[\frac{\partial \theta}{\partial \rho_{12}}\right] = \mathrm{sign}\left[\frac{\partial \mathbf{E}[r]}{\partial \rho_{12}}\right] > 0$. Similarly, if $A_{11} - A_{12} < 0$ and $A_{22} - A_{21} < 0$, then $\mathrm{sign}\left[\frac{\partial \theta}{\partial \rho_{12}}\right] = \mathrm{sign}\left[\frac{\partial \mathbf{E}[r]}{\partial \rho_{12}}\right] < 0$. This result completes the proof.

$\square$

Appendix B.   Portfolio Composition

Which pair strategies tend to be active in the pairs-based portfolio, Pairs($mom = 12, win = 120, n = 4, p = 3$), of Section 4? Exhibit B.1 reports the top 12 active pair strategies by frequency of inclusion in the top $n = 4$ pairs over the evaluation period. Short-Term Bonds are a frequent pair-strategy asset partner among top active pairs and tend to pair with traditionally risky assets such as high-yield bonds, stocks, and commodities. Through the lens of our theory, this phenomenon is not surprising given that the returns of Short-Term Bonds are relatively persistent over time (own-asset predictability) and are negatively correlated with the returns of risky assets. These properties increase θ, thereby increasing the frequency of the appearance of Short-Term Bonds among top pair strategies.

Note, however, that the frequent pairing of Short-Term Bonds and risky assets is not an artifact of the relative means or volatilities of these assets. First, we demean and volatility scale all raw signals to have unit variance. Second, θ is entirely driven by scale-independent measures of the dependence among signals and returns, namely, signal-signal and signal-return *correlations*. Therefore, the marginal distributions of signals and returns have no bearing on pair-strategy rankings.

Other frequently active pairs include Global Core Bonds and EM Equities, Long Treasury Bonds and EM Equities, Short-Term Bonds and Long Treasury Bonds, and High Yield and US Equities. These pairings arise endogenously as a function of scale-independent measures of joint dependence.

24

<div align="center">

Exhibit B.1: Frequently Active Pair Strategies

</div>

| Pair Strategy Rank (High to Low Freq.)	Assets in Pair Strategy	Freq. of Pair in Top $n=4$ (%)		Asset Legend	
1	1, 6	62.7		1	Short-Term Bonds
2	1, 12	59.4		2	US Core Bonds
3	1, 13	59.0		3	Long Treasury Bonds
4	1, 11	53.7		4	Long Credit Bonds
5	1, 10	40.6		5	Global Core Bonds
6	1, 9	36.1		6	High Yield
7	5, 13	31.1		7	EM USD Bonds
8	1, 8	21.7		8	Commodities
9	3, 13	14.8		9	REITs
10	1, 3	4.1		10	US Equities Small
11	6, 11	3.3		11	US Equities
12	3, 8	3.3		12	Dev ex US Equities
				13	EM Equities

Notes: This exhibit reports the top 12 pair strategies of Pairs($mom = 12, win = 120, n = 4$) by their frequency of appearance in the top $n = 4$ pairs over the evaluation period 2000-01 to 2020-05.

Exhibit B.2 plots the time series of asset positions under the Pairs strategy. Asset positions are relatively stable over the evaluation period, which reflects the relative stability of the θ's and explains the relatively low turnover of the strategy. Leading into the major crisis subperiods (2008-2009 and 2020-01 to 2020-05), risky assets move from sustained phases of overweighting to being underweighted, while safe assets move from sustained phases of underweighting to being overweighted. These types of dynamics help explain the favorable performance of the strategy overall and during crisis subperiods.

Appendix C.   Robustness

In this section, we analyze the robustness of our running example pairs-based portfolio, Pairs($mom = 12, win = 120, n = 4, p = 3$), along several dimensions. First, the pairs-based portfolio exhibits consistent performance in subsamples. Exhibit C.1 shows the 10-year rolling Sharpe ratio time series for both Pairs and CSMOM portfolios. Both series exhibit a downward trend in performance over the evaluation period, which is consistent with the weakening of performance for momentum portfolios in the recent decade (Babu et al., 2020 and Garg et al., 2019, 2020 offer explanations for this broader phenomenon). However, the Pairs series level exceeds that of CSMOM uniformly over the sample and exhibits less deterioration. CSMOM falls from a Sharpe ratio of about 0.50 to 0.00 by the end of the

Exhibit B.2: Time Series of Asset Positions
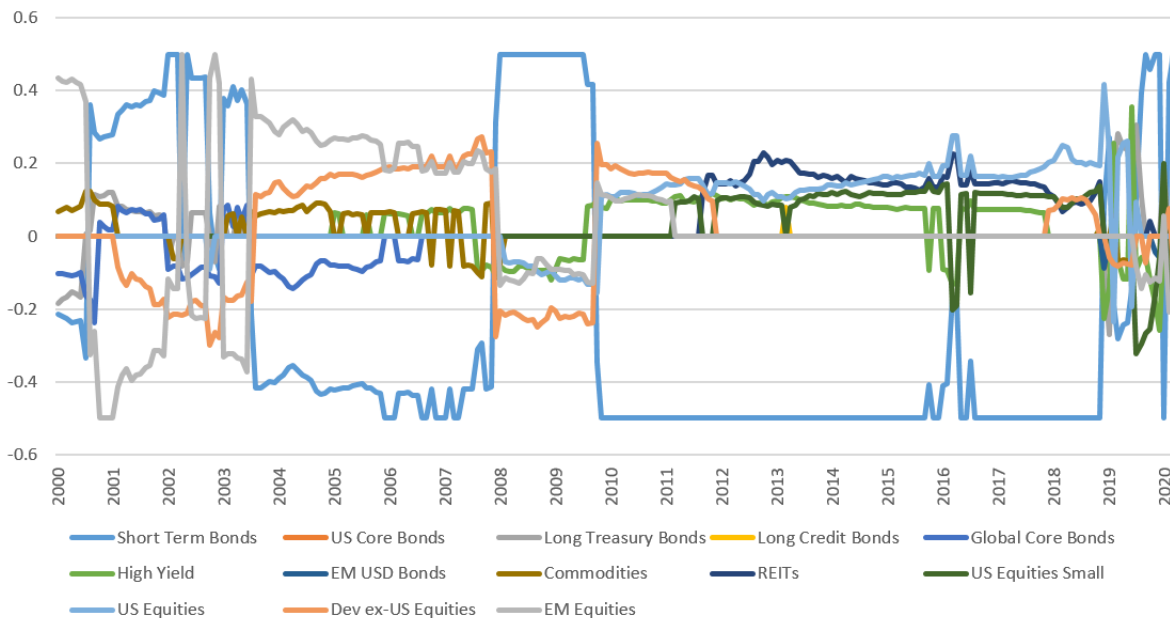
Notes: This exhibit reports the time series of asset positions in the Pairs($mom = 12, win = 120, n = 4, p = 3$) portfolio over the evaluation period 2000-01 to 2020-05.

sample, whereas Pairs starts at about 0.80 and ends the sample at about 0.60.

Second, the pairs-based portfolio exhibits consistent performance across momentum look-back horizons. Exhibit C.2 shows that the Sharpe ratio of Pairs for different values of mom ranging from 1 to 12 months are all within the range of 0.47 to 0.69, with all except $mom = 4$ above 0.55.

Third, the performance surface of Pairs is relatively smooth across values of n and p. Exhibit C.3 reports the annualized average return and Sharpe ratio of Pairs for an array of values for the number of top pairs n and tilting parameter p. Exhibit C.3 shows that the values for our running example (highlighted) are not the best values of n and p over the evaluation period. Nevertheless, it also shows that deviations of n and p have a mild impact on performance. Even for higher values of n and low values of p, the performance of Pairs is well above that of CSMOM. Moreover, lower values of n can improve average returns in the historical simulation without adding relatively much volatility, as evidenced by the Sharpe ratios for lower values of n.

Fourth, we analyze robustness to the estimation window parameter win. Exhibit C.4 reports the annualized Sharpe ratio and number of active pairs for several momentum horizons at various estimation window lengths ranging from 1 year to 30 years. For estimation

26

Exhibit C.1: 10-Year Rolling Sharpe Ratios
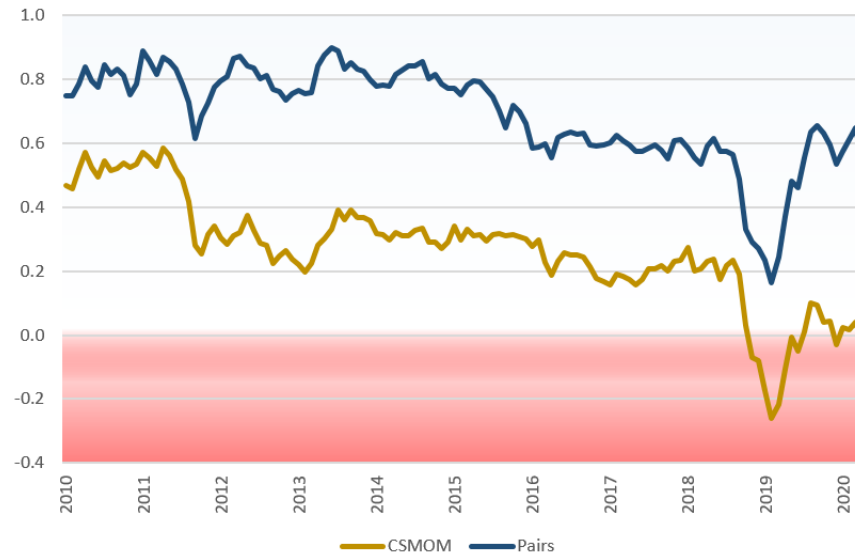
Notes: This exhibit plots the time series of 10-year rolling Sharpe ratios (anlzd.) for CSMOM($mom = 12, LNS = 1/3$) and Pairs($mom = 12, win = 120, n = 4, p = 3$) over the evaluation period 2000-01 to 2020-05.

Exhibit C.2: Performance Across Momentum Horizons
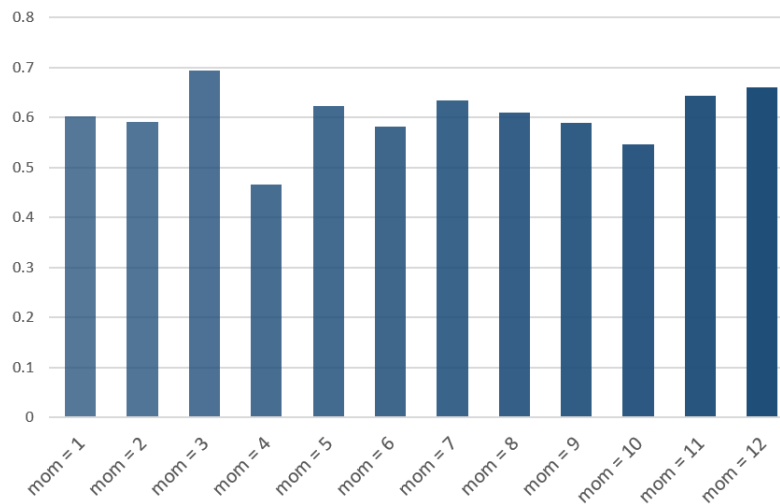
Notes: This exhibit reports the Sharpe ratios (anlzd.) of Pairs($mom = k, win = 120, n = 4, p = 3$) for $k = 1, 2, \ldots, 12$ over the evaluation period 2000-01 to 2020-05.

windows below 60 months, performance is weaker and noisier for many momentum horizons. Also, the number of active pairs is high, indicating that the θ's of the pairs reflect noise rather

27

Exhibit C.3: Performance Across n and p

	(a)						(b)			
	Average Return (anlzd.)						Sharpe Ratio (anlzd.)			
n	$p=0$	$p=1$	$p=3$	$p=5$		n	$p=0$	$p=1$	$p=3$	$p=5$
1	6.0	6.0	6.0	6.0		1	0.64	0.64	0.64	0.64
2	6.4	6.5	6.5	6.5		2	0.75	0.75	0.75	0.74
3	5.2	5.4	5.7	5.9		3	0.62	0.64	0.67	0.69
4	4.8	5.0	5.4	5.8		4	0.60	0.63	0.67	0.69
5	4.4	4.7	5.2	5.6		5	0.56	0.60	0.65	0.68
6	4.2	4.6	5.2	5.6		6	0.56	0.60	0.66	0.69
7	4.0	4.4	5.1	5.6		7	0.54	0.59	0.65	0.69
8	3.7	4.1	4.9	5.5		8	0.51	0.56	0.64	0.68
9	3.7	4.1	4.9	5.5		9	0.52	0.56	0.64	0.68
10	3.4	3.9	4.9	5.5		10	0.49	0.55	0.64	0.68
15	3.7	4.2	5.0	5.5		15	0.55	0.60	0.66	0.69
20	3.6	4.1	5.0	5.5		20	0.55	0.60	0.66	0.69

Notes: This exhibit reports (a) the average return (anlzd.) and (b) Sharpe ratio (anlzd.) of Pairs($mom = 12, win = 120, n, p$) for various values of n and p over the evaluation period 2000-01 to 2020-04. The values for our running example ($n = 4, p = 3$) are highlighted.

than fundamentals. For estimation windows above 180 months, both the performance and number of active pairs are relatively insensitive to the estimation window. For estimation windows in the middle of the range, centered around 120 months, performance appears to balance the tradeoff between the noise that comes with shorter estimation windows and the staleness that comes with longer windows. Likewise, that balance is reflected in the number of active pairs for the middle range of estimation windows.

Exhibit C.4: Sharpe Ratios and Number of Active Pairs Across Estimation Windows for Various Momentum Horizons
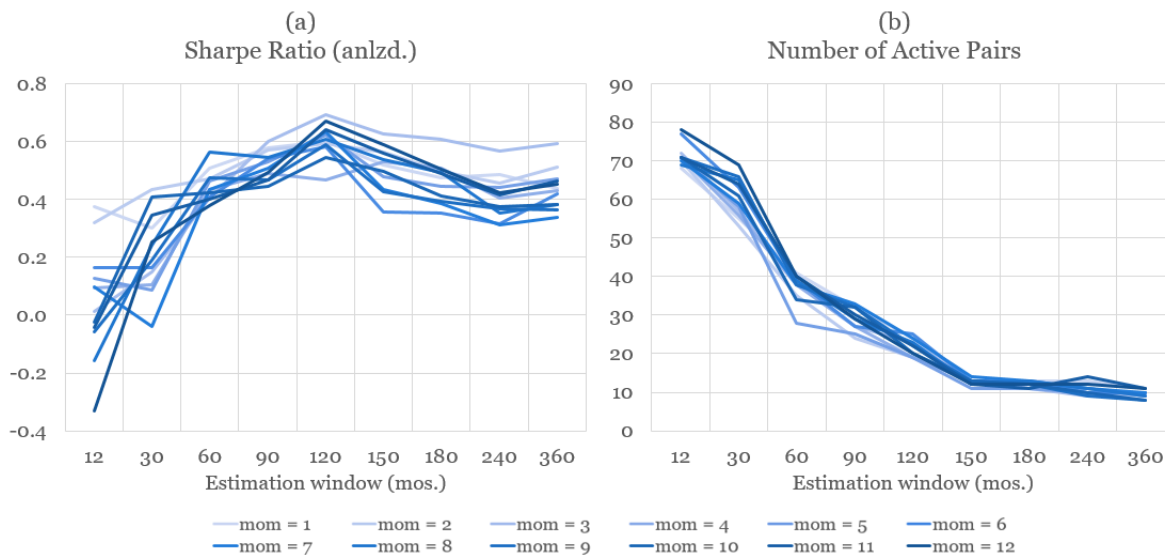
Notes: This exhibit reports (a) the Sharpe ratio (anlzd.) and (b) the number of active pairs of Pairs($mom, win, n = 4, p = 3$) for various values of win ranging from 12 to 360 months at 30-month intermediate step sizes for various values of $mom = 1, 2, \ldots, 12$ over the evaluation period 2000-01 to 2020-05.

References

Andersson, M., Krylova, E., Vähämaa, S., 2008. Why does the correlation between stock and bond returns vary over time? Applied Financial Economics 18, 139–151.

Asness, C. S., 1994. Variables that explain stock returns. Ph.D. thesis, University of Chicago.

Asness, C. S., Moskowitz, T. J., Pedersen, L. H., 2013. Value and momentum everywhere. Journal of Finance 68, 929–985.

Babu, A., Hoffman, B., Levine, A., Ooi, Y. H., Schroeder, S., Stamelos, E., 2020. You can't always trend when you want. Journal of Portfolio Management 46, 1–17.

Baz, J., Granger, N., Harvey, C. R., Roux, N. L., Rattray, S., 2015. Dissecting investment strategies in the cross section and time series, available at SSRN 2695101.

Beltratti, A. E., Shiller, R. J., 1992. Stock prices and bond yields. Journal of Monetary Economics 30, 25–46.

Boucrelle, C., Fur, Y. L., Solnik, B., 1996. International market correlation and volatility. Financial Analysts Journal 52, 17–34.

Campbell, R., Koedijk, K., Kofman, P., 2002. Increased correlation in bear markets. Financial Analysts Journal 58, 87.

Cappiello, L., Engle, R. F., Sheppard, K., 2006. Asymmetric dynamics in the correlations of global equity and bond returns. Journal of Financial Econometrics 4, 537–572.

Chen, S. P., He, L. Y., 2001. Multifractal detrended cross-correlation analysis of agricultural futures markets. Chaos, Solitons and Fractals 44, 355–361.

Conrad, J., Kaul, G., 1998. An anatomy of trading strategies. Review of Financial Studies 11, 489–519.

Engle, R. F., Ito, T., Lin, W. L., 1994. Do bulls and bears move across borders? international transmission of stock returns and volatility. Review of Financial Studies 7, 507–538.

Erb, C. B., Harvey, C. R., 2006. The tactical and strategic value of commodity futures. Financial Analysts Journal 62, 69–97.

Erb, C. B., Harvey, C. R., Viskanta, T., 1994. Forecasting international correlation. Financial Analysts Journal 50, 32–45.

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., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Farago, A., Hjalmarsson, E., 2019. Stock price co-movement and the foundations of pairs trading. Journal of Financial and Quantitative Analysis 54, 629–665.

Garg, A., Goulding, C. L., Harvey, C. R., Mazzoleni, M. G., 2019. Momentum turning points, available at SSRN 3489539.

Garg, A., Goulding, C. L., Harvey, C. R., Mazzoleni, M. G., 2020. Breaking bad trends, available at SSRN 3594888.

Gatev, E., Goetzmann, W. N., Rouwenhorst, K. G., 2006. Pairs trading: Performance of a relative-value arbitrage rule. Review of Financial Studies 19, 797–827.

Griffin, J. M., Ji, S., Martin, J. S., 2003. Momentum investing and business cycle risk: Evidence from pole to pole. Journal of Finance 58, 2515–2547.

Gulko, L., 2002. Decoupling. Journal of Portfolio Management 28, 59–67.

Gutierrez, R. C., Kelley, E. K., 2008. The long-lasting momentum in weekly returns. Journal of Finance 63, 415–447.

Harvey, C. R., 2017. Presidential address: The scientific outlook in financial economics. Journal of Finance 72, 1399–1440.

Harvey, C. R., Liu, Y., Zhu, H., 2016. ...and the cross-section of expected returns. Review of Financial Studies 29, 5–68.

Israel, R., Moskowitz, J. T., 2012. The role of shorting, firm size, and time on market anomalies. Journal of Financial Economics 108, 275–301.

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., 2001. Profitability of momentum strategies: An evaluation of alternative explanations. Journal of Finance 56, 699–720.

Karolyi, G. A., Stulz, R. M., 1996. Why do markets move together? Journal of Finance 51, 951–986.

Kelly, B. T., Malamud, S., Pedersen, L. H., 2020. Principal portoflios, available at SSRN 3623983.

Koijen, R. S. J., Moskowitz, T. J., Pedersen, L. H., Vrugt, E. B., 2018. Carry. Journal of Financial Economics 127, 197–225.

Lee, C., Swaminathan, B., 2000. Price momentum and trading volume. Journal of Finance 55, 2017–2069.

Longin, F., Solnik, B., 1995. Is the correlation in international equity returns constant? Journal of International Money and Finance 14, 3–26.

Maddala, G. S., 1983. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press.

Markowitz, H., 1952. Portfolio selection. Journal of Finance 7, 77–91.

Mizuno, T., Takayasu, H., Takayasu, M., 2006. Correlation networks among currencies. Physica A: Statistical Mechanics and its Applications 364, 336–342.

Ramchand, L., Susmel, R., 1998. Volatility and cross correlation among major stock markets. Journal of Empirical Finance 5, 397–416.

Rouwenhorst, K. G., 1998. International momentum strategies. Journal of Finance 53, 267–284.

Roy, A., 1952. Safety first and the holding of assets. Econometrica 20, 431–449.

Santis, G. D., Gerard, B., 1997. International asset pricing and portfolio diversification with time-varying risk. Journal of Finance 52, 1881–1912.

Vidyamurthy, G., 2004. Pairs Trading: Quantitative Methods and Analysis. Wiley, New York.

Whistler, M., 2004. Trading Pairs: Capturing Profits and Hedging Risk with Statistical Arbitrage Strategies. Wiley, New York.