Internet Appendix for "Comomentum: Inferring Arbitrage Activity from Return Correlations"

Dong Lou and Christopher Polk

London School of Economics

This version: October 2021

1 A Simple Model

To help motivate and interpret our empirical work, we develop a simple model of crowded trading, building on the work of Hong and Stein (1999). In their model, Hong and Stein study the interaction of newswatchers and quantitative traders following well-known strategies, and assume that the number of quantitative traders is fixed. We extend their setting to incorporate a *time-varying* number of quantitative traders (or, equivalently, time-varying aggregate risk tolerance), to reflect the fact that arbitrage capital can change stochastically.

1.1 Baseline Model Setup

Following Hong and Stein, our economy has a single asset with supply Q that pays a liquidating dividend, D_T, at time T. Specifically,

$$D_T = D_0 + \sum_{j=0}^{T} \varepsilon_j \tag{1}$$

where ε_j is iid $N(0, \sigma^2)$. All ε's are independently distributed and can be further decomposed into Z independent sub-innovations, $\varepsilon_j = \varepsilon_j^1 + ... + \varepsilon_j^z$, where ε_j^l is iid $N(0, \sigma^2/z)$. Thus, each sub-innovation has the same variance. Throughout, we consider the limiting case where T goes to infinity. This choice allows us to focus on the steady-state trading strategy—that is, a strategy that does not depend on how close we are to the terminal payoff.

Our baseline model features three groups of investors. The first group is newswatchers with CARA utility. These newswatchers are divided into z equal-sized groups. At time t, news about ε_{t+z-1} starts to spread. More specifically, at time t, group 1 newswatchers observe ε_{t+z-1}^1, group 2 observes ε_{t+z-1}^2, and so forth. Thus, at time t, each sub-innovation of ε_{t+z-1} is observed by 1/z of the total newswatcher population. Next, at time $t+1$, the groups rotate, so that group 1 newswatchers observe ε_{t+z-1}^2, group 2 observes ε_{t+z-1}^3, and so forth. Therefore, at time $t+1$, each sub-innovation of of ε_{t+z-1} is observed by 2/z of the total newswatcher population. This rotation process continues until time $t+z-1$, at which point each group of newswatchers has observed the entire signal ε_{t+z-1}, so that the entirety of ε_{t+z-1} is public information. Importantly, newswatchers do not condition their demand on P_t.[1] As shown by Hong and Stein, this structure naturally leads to underreaction:

$$P_t = D_t + \{(z-1)\varepsilon_{t+1} + (z-2)\varepsilon_{t+2} + ... + \varepsilon_{t+z-1}\}/z - \theta Q \tag{2}$$

Note that θ is a function of newswatchers' risk aversion and σ^2. For simplicity, we normalize θ to 1.

Our second group of investors is momentum traders, who also have CARA utility with risk tolerance of γ^{MOM}. Following Hong and Stein (1999), we assume that these traders hold their positions for j^{MOM} periods, and, importantly, only condition on $\Delta P_{t-1} = P_{t-1} - P_{t-2}$. In other words, generation-t momentum traders' demand is $F_t^{MOM} = \phi_t^{MOM} \Delta P_{t-1}$, where ϕ_t^{MOM} is the endogenous response chosen by momentum traders.

Our contribution relative to Hong and Stein is that we allow the number of momentum traders to be time varying as captured by N_t^{MOM}. In particular, we assume that N_t^{MOM} is the sum of two components, one observable (X_t^{MOM}) and one unobservable (Y_t^{MOM}). The evolution of X_t^{MOM} follows a two-state ($0.5 - Spread_X$, $0.5 + Spread_X$) Markov process where the persistence of each state is calibrated to the AR(1) coefficient (0.66) of our comomentum measure in the data. Y_t^{MOM} also follows a two-state ($0.5 - Spread_Y$, $0.5 + Spread_Y$) Markov process with the same persistence.[2] In the special case where $Spread_Y = 0$, the total number of momentum traders is effectively publicly observable; otherwise, momentum traders only see a noisy proxy.

Our final group of investors consists of value traders, also having CARA utility with risk tolerance of γ^{VAL}. Value traders hold their positions for j^{VAL} periods, and condition on $\hat{P}_t - P_t$, where $\hat{P}_t$ is a noisy signal of the fundamental value of the firm (i.e., firm value incorporating all available information at t plus an independent noise term, $\varepsilon_t^{\hat{P}} \sim N(0, \sigma_{\hat{P}}^2)$). Thus, $\hat{P}_t - P_t$ is reminiscent of the value signal used in practice. Consequently, generation-t value traders' demand is $F_t^{VAL} = \phi_t^{VAL}(\hat{P}_t - P_t)$, where ϕ_t^{VAL} is the endogenous response chosen by value traders. A key distinction between the momentum and value strategies is that while momentum trading exacerbates the initial momentum signal, value traders shrink the value signal; in other words, the former is a positive-feedback strategy and the latter a negative-feedback strategy.

[1] Hong and Stein (1999) provide extensive discussion of the way in which their model departs from a classic model like Grossman (1976). They also show how the important takeaways from their theory continue to hold even if they add fully-rational but risk-averse arbitrageurs who can condition on everything in the model that is observed by any other trader.

[2] We require that $Spread_X < 0.5$ and $Spread_Y < 0.5$, so that the amount of arbitrage capital is strictly positive in each period.

We further assume that N_t^{VAL}, the number of value traders in period t, follows a similar structure to N_t^{MOM} (with two components, both of which follow a two-state Markov process). For simplicity, we assume that X_t^{MOM} is observable only for momentum traders, while X_t^{VAL} is observable only for value traders. (We relax this assumption in an extension where a single group of arbitrageurs optimally trade both signals.) Given these assumptions, the aggregate supply absorbed by newswatchers is thus

$$\begin{aligned} S_t &= Q - \sum_{i=1}^{j^{MOM}} F_{t-i+1}^{MOM} - \sum_{i=1}^{j^{VAL}} F_{t-i+1}^{VAL} \\ &= Q - \sum_{i=1}^{j^{MOM}} N_{t-i+1}^{MOM} \phi_{t-i+1}^{MOM} \Delta P_{t-i} - \sum_{i=1}^{j^{VAL}} N_{t-i+1}^{VAL} \phi_{t-i+1}^{VAL} (\hat{P}_{t-i+1} - P_{t-i+1}). \end{aligned} \tag{3}$$

1.2 Equilibrium

We conjecture a linear pricing equation

$$\begin{aligned} P_t &= D_t + ((z-1)\varepsilon_{t+1} + (z-2)\varepsilon_{t+2} + ... + \varepsilon_{t+z-1})/z \\ &\quad -Q + \sum_{i=1}^{j^{MOM}} N_{t-i+1}^{MOM} \phi_{t-i+1}^{MOM} \Delta P_{t-i} + \sum_{i=1}^{j^{VAL}} N_{t-i+1}^{VAL} \phi_{t-i+1}^{VAL} (\hat{P}_{t-i+1} - P_{t-i+1}). \end{aligned} \tag{4}$$

Without loss of generality and for the purpose of illustration, we set $j^{MOM} = j^{VAL} = 1$ (i.e., momentum traders and value traders live for only one period).[3] As a consequence, demand by momentum and value traders can be characterized by the following equations:

$$\begin{aligned} \phi_t^{MOM}(X_t^{MOM}, X_t^{VAL})\Delta P_{t-1} &= \frac{E_{MOM}(\Delta P_{t+1}|X_t^{MOM}, X_t^{VAL})}{Var(\Delta P_{t+1}|X_t^{MOM}, X_t^{VAL})}\gamma^{MOM} \\ \phi_t^{VAL}(X_t^{MOM}, X_t^{VAL})(\hat{P}_t - P_t) &= \frac{E_{VAL}(\Delta P_{t+1}|X_t^{MOM}, X_t^{VAL})}{Var(\Delta P_{t+1}|X_t^{MOM}, X_t^{VAL})}\gamma^{VAL} \end{aligned}$$

[3] In other words, all momentum and value traders are reborn in each period. Our theoretical conclusions are robust to allowing j to differ from one.

where the expected next-period return can be expressed as a linear projection onto the momentum and value signals:

$$E^{MOM}(\Delta P_{t+1}|\Delta P_{t-1}, X_t^{MOM}, X_t^{VAL}) = \frac{Cov(\Delta P_{t+1}, \Delta P_{t-1}|X_t^{MOM}, X_t^{VAL})}{Var(\Delta P_{t-1}|X_t^{MOM}, X_t^{VAL})}\Delta P_{t-1}$$

$$E^{VAL}(\Delta P_{t+1}|\hat{P}_t - P_t, X_t^{MOM}, X_t^{VAL}) = \frac{Cov(\Delta P_{t+1}, \hat{P}_t - P_t|X_t^{MOM}, X_t^{VAL})}{Var(\hat{P}_t - P_t|X_t^{MOM}, X_t^{VAL})}(\hat{P}_t - P_t)$$

1.3 An Extension

We next extend the model by allowing N_t quantitative traders to use an optimal blended strategy. In other words, every quantitative trader now runs a bivariate regression to forecast future stock returns based on both the momentum and value signals. As a consequence, demand by arbitrageurs can be characterized by the following:

$$\phi_t^{MOM}\Delta P_{t-1} + \phi_t^{VAL}(\hat{P}_t - P_t) = \frac{E(\Delta P_{t+1}|X_t)}{Var(\Delta P_{t+1}|X_t)}\gamma, \tag{5}$$

where γ is the risk tolerance of all quantitative traders,and X_t is the observable variation in arbitrage capital. The expected next-period return is now

$$E(\Delta P_{t+1}|X_t, \Delta P_{t-1}, \hat{P}_t - P_t) = \begin{bmatrix}\Delta P_{t-1} & \hat{P}_t - P_t\end{bmatrix} \Sigma^{-1} \begin{bmatrix}Cov(\Delta P_{t+1}, \Delta P_{t-1}|X_t) \\ Cov(\Delta P_{t+1}, \hat{P}_t - P_t|X_t)\end{bmatrix},$$

where

$$\Sigma = \begin{bmatrix} Var(\Delta P_{t-1}|X_t) & Cov(\Delta P_{t-1}, \hat{P}_t - P_t|X_t) \\ Cov(\hat{P}_t - P_t, \Delta P_{t-1}|X_t) & Var(\hat{P}_t - P_t|X_t) \end{bmatrix}.$$

We can then rewrite ϕ_t^{MOM} and ϕ_t^{VAL} as

$$\begin{bmatrix}\phi_t^{MOM}(X_t) \\ \phi_t^{VAL}(X_t)\end{bmatrix} = \Sigma^{-1} \begin{bmatrix}Cov(\Delta P_{t+1}, \Delta P_{t-1}|X_t) \\ Cov(\Delta P_{t+1}, \hat{P}_t - P_t|X_t)\end{bmatrix} \frac{\gamma}{Var(\Delta P_{t+1}|X_t)}$$

and conjecture a linear pricing equation:

$$\begin{aligned} P_t &= D_t + ((z-1)\varepsilon_{t+1} + (z-2)\varepsilon_{t+2} + ... + \varepsilon_{t+z-1})/z \\ &\quad -Q + N_t[\phi_t^{MOM}\Delta P_{t-1} + \phi_t^{VAL}(\hat{P}_t - P_t)]. \end{aligned} \tag{6}$$

1.4 Model Solution

These equations do not have an analytical solution for ϕ_t^{MOM} and ϕ_t^{VAL}, which are functions of state variables X_t^{MOM} and X_t^{VAL}. Thus, we solve the model via simulation. Specifically, we set the initial values of ϕ_t^{MOM} and ϕ_t^{VAL} to the solutions based on constant N_t^{MOM} and N_t^{VAL} (as in the Hong and Stein case). In each subsequent round, we simulate a price path based on ϕ_t^{MOM} and ϕ_t^{VAL} computed in the previous iteration; we then re-calculate ϕ_t^{MOM} and ϕ_t^{VAL} using the simulated price path. The solution to the above set of equations is then the fixed point in ϕ_t^{MOM} and ϕ_t^{VAL} conditional on X_t^{MOM} and X_t^{VAL}.

Section 2.1 of the Internet Appendix explores a variety of implications of our model via comparative statics. We summarize the key results as follows. First, momentum returns peak in the short run and then gradually and partially reverse in subsequent periods. Second, all else equal, a larger amount of momentum capital (i.e., when either X_t^{MOM} or Y_t^{MOM} is relatively high) is associated with a larger return effect at the time of the arbitrageurs' trades and then a smaller drift subsequently. In other words, as more capital arrives, momentum traders incorporate more information into prices as they trade, consequently making the momentum strategy less profitable. Third, a larger amount of momentum capital is also associated with a larger reversal in the long run, consistent with the idea that momentum trading can be destabilizing.

1.5 Arbitrage Activity and Strategy Returns

To speak more directly to our empirical tests, we compare periods with high momentum (value) activity vs. periods with low momentum (value) activity in our simulated data. As shown in Equation (3), arbitrage demand for the momentum strategy in our simple model is $N_t^{MOM}\phi_t^{MOM}abs(\Delta P_{t-1})$ (i.e., the number of momentum traders times their momentum trading intensity times the magnitude of the momentum signal), and that for value is $N_t^{VAL}\phi_t^{VAL}abs(\hat{P}_t - P_t)$ (i.e., the number of value traders times their value trading intensity times the magnitude of the value signal). Consequently, we sort all simulated periods into two groups based on the absolute value of either $N_t^{MOM}\phi_t^{MOM}\Delta P_{t-1}$ or $N_t^{VAL}\phi_t^{VAL}(\hat{P}_t - P_t)$ and track the returns to the momentum and value strategies in subsequent periods. Note that if the momentum (or value) signal is negative, arbitrageurs short the stock, so their profit is minus the stock return when that occurs.

The simulation results are shown in Appendix Table A1. Panel A presents momentum returns after periods of high and low momentum demand. Specifically, we report momentum returns in period 0 (when momentum traders put on the trade), in period 1 (when momentum traders hold their positions), and periods 2-12 (the long-run). Across all γ^{MOM} values, momentum spreads in the formation period are larger in high momentum activity periods than in low activity periods. There is also a larger reversal to the momentum strategy after periods with high momentum activity than periods with low activity, again across all γ^{MOM} values. Momentum returns in the holding period (j=1 in this case), however, depend on the γ^{MOM} value. When momentum traders are relatively risk-averse (i.e., with a low γ^{MOM}), momentum returns in period 1 are larger in high momentum activity periods than in low activity periods; when momentum traders are relatively risk-tolerant (i.e., with a high γ^{MOM}), the reverse is true.

Panel B of the same table presents value returns after periods of high and low value strategy activity. In both the holding and post-holding periods, value returns are strictly larger after high realizations of value activity than low realizations of value activity. This is true for all γ^{VAL} values. The reason that our model has different predictions for momentum and value is that value is a negative-feedback strategy. Though value traders' demand is a linear function of the value signal as well, the resulting demand also shrinks the signal. This fact has two implications. 1) The equilibrium value signal is a positive predictor of future value strategy returns (in contrast, the momentum signal can be a negative predictor of future returns). 2) High value activity is accompanied by a relatively large value signal, which in turn indicates high expected returns to the value strategy. We provide a more detailed discussion of momentum vs. value strategies in Section 2.2 of this appendix.

2 Numerical Comparative Statics

2.1 Momentum Strategy Returns

We conduct a number of numerical comparative-static exercises for momentum strategy returns using the model described above. For each set of parameter values, we calculate the cumulative momentum profits, i.e., the coefficient estimates from the regression of ΔP_{t+k} on ΔP_{t-1}, where k ranges from 1 to 12. Period 0 is when arbitrageurs put on their trades and period 1 is when momentum profits are realized. For illustration purposes, we focus on the setting where z is equal to 3 (i.e., the signal in each period is divided into three sub-signals); we have also experimented with other z values and obtained similar patterns (results available upon request). When unstated, γ^{MOM} is set to 2 and γ^{VAL} to 1 in the baseline model, and γ to 3 in the extended model; for other parameters, the default values of $Spread_X$ and $Spread_Y$ are set to 0.4 and $\sigma_{\hat{P}}$ to 1.

Appendix Figures A1-A4 present comparative statics of the baseline model, in which one group of arbitrageurs follows a momentum strategy and the other group follows a value strategy (each solving a univariate optimization problem). Each graph in these figures contains three to five lines that correspond to different realizations of X_t^{MOM} and Y_t^{MOM}. "High" or "Low" denotes that the realization of X_t^{MOM} or Y_t^{MOM} is either high or low. "Unconditional" or "Uncond" corresponds to settings in which the cumulative momentum return is unconditional on X_t^{MOM} or Y_t^{MOM}, or both.

A number of general observations are worth pointing out. First, momentum returns peak in period 1 (or 0, depending on the parameter values), and gradually reverse in the subsequent periods. As evident by the positive cumulative return at the end of period 12, momentum returns are only partially reversed. Second, all else equal, a larger amount of momentum capital (i.e., in the high state of X_t^{MOM} or Y_t^{MOM}) is associated with a larger return effect in period 0 (when arbitrageurs put on their trades) and a smaller drift in period 1; in other words, as more capital arrives, momentum traders incorporate more information into prices as they trade, consequently making the momentum strategy less profitable. Third, a larger amount of momentum capital is also associated with a larger reversal in the long run (periods 2 to 12), consistent with the idea that momentum trading can be destabilizing.

Appendix Figure A1 depicts cumulative momentum returns as a function of the risk tolerance of momentum traders, with γ^{MOM} ranging from 1 to 10. Consistent with Hong and Stein (1999), we find that across all settings, as momentum traders become more risk tolerant, a) the average momentum profit in period 1 (as well as the cumulative return over the entire 12 periods) declines, and b) the long-run reversal to the momentum strategy in periods 2 to 12 increases. Put differently, as arbitrageurs trade more aggressively on the momentum signal, they reduce the short-term profitability of the strategy and at the same time induce a stronger long-run reversal.

Appendix Figure A2 shows momentum returns with varying degrees of risk tolerance of value traders, where γ^{VAL} ranges from 1 to 10. Since value traders observe a noisy signal of the fundamental value, they trade to correct any deviation from this benchmark value – that is, value investors temper whatever price effect caused by momentum traders. Not surprisingly, as γ^{VAL} increases, the entire curve of cumulative momentum returns shifts downward: a) a smaller momentum spread in period 0, b) a smaller momentum return in period 1, and c) a weaker reversal in periods 2 to 12.

In Appendix Figure A3, we keep the two risk-tolerance coefficients constant, and vary the size of $Spread_X$, the part of time-varying momentum capital observable to momentum traders. As $Spread_X$ increases from 0.05 to 0.5, the unconditional momentum profit remains roughly constant, but the wedge between the high and low states of X_t^{MOM} is magnified. For example, when $Spread_X$ is close to zero, the three lines (high X_t^{MOM}, unconditional, low X_t^{MOM}) are virtually indistinguishable from each other. In contrast, in the bottom-right graph, where $Spread_X$=0.5, the high-X_t^{MOM} line stays visibly above the low-X_t^{MOM} line. (This is the case even if we set Y_t^{MOM} to zero.)

This result may seem surprising initially, as momentum traders observe the realizations of X_t^{MOM} and adjust their trading intensity accordingly. While this intuition is correct, momentum traders' endogenous response (captured by ϕ_t^{MOM}) does not fully offset the variation in X_t^{MOM}. This arises because momentum traders are risk averse; as such, they demand a higher premium to bet more aggressively on momentum in the low-X_t^{MOM} state. In other words, the product of N_t^{MOM} and ϕ_t^{MOM} remains higher in the high-X_t^{MOM} state, despite the fact that ϕ_t^{MOM} is higher in the low-X_t^{MOM} state. The table below documents this fact by showing both ϕ_t and $N_t \times \phi_t$ in the high and low X states as we vary $Spread_X$.

ϕ in the Baseline Model				
	ϕ^{MOM}		ϕ^{VAL}	
$Spread_X$	High	Low	High	Low
0.05	0.335	0.359	0.266	0.279
0.1	0.323	0.371	0.260	0.285
0.2	0.302	0.397	0.249	0.298
0.3	0.284	0.429	0.239	0.313
0.4	0.268	0.467	0.229	0.328
0.5	0.253	0.513	0.219	0.345
	$N^{MOM}\phi^{MOM}$		$N^{VAL}\phi^{VAL}$	
$Spread_X$	High	Low	High	Low
0.05	0.351	0.341	0.279	0.265
0.1	0.355	0.334	0.286	0.256
0.2	0.362	0.318	0.299	0.238
0.3	0.370	0.300	0.310	0.219
0.4	0.375	0.280	0.320	0.197
0.5	0.379	0.257	0.328	0.172

Appendix Figure A4 repeats the exercise in Figure A3 by introducing variation in $Spread_Y$, the unobservable component of time-varying momentum capital. The general pattern is similar to that in Figure A3: as we increase $Spread_Y$, the wedge in short-run momentum returns and long-run return reversal between the high and low states of Y_t^{MOM} is magnified. Not surprisingly, all else equal, the effect of $Spread_Y$ on the wedge is larger than that of $Spread_X$. This is because momentum traders do not observe Y_t^{MOM}, and therefore cannot condition their trading intensity on the realizations of Y_t^{MOM}. Put another way, ϕ_t^{MOM} does not vary across the high vs. low Y_t^{MOM} states, so the product of N_t^{MOM} and ϕ_t^{MOM} is solely determined by the realization of Y_t^{MOM}.

Appendix Figures A5-A7 present comparative statics of the extended model, in which a single group of arbitrageurs, whose time-varying capital is denoted by N_t, trade both the momentum and value strategies (i.e., solving a bivariate optimization problem).

Appendix Figure A5 shows cumulative momentum returns as a function of the risk tolerance of this single group of arbitrageurs. As γ increases (from 1 to 10), arbitrageurs bet more aggressively on both momentum and value (the latter again helps mitigate any price effect induced by the former). Consequently, returns in both period 0 (when arbitrageurs put their trades) and period 1 (when momentum profits are realized) decrease; the long-run reversal of momentum also gets weaker. Moreover, the cumulative return at the end of period 12 (reflecting the overall underreaction to news) gets smaller as a result of intensified trading by arbitrageurs.

Appendix Figures A6 and A7 examine momentum returns with different values of $Spread_X$ and $Spread_Y$, respectively. (The table below shows ϕ_t and $N_t \times \phi_t$ in the high and low X states with different values of $Spread_X$.) The patterns are similar to those in Appendix Figures A3 and A4. Again, an increase in *Spread* (in either X_t or Y_t) leads to a larger wedge in short-term momentum profits as well as in long-term return reversal across the high and low states of the corresponding component in momentum capital. Moreover, since arbitrageurs can only observe X_t but not Y_t, $Spread_Y$ has a larger impact on momentum returns than $Spread_X$.[4]

ϕ in the Extended Model

	ϕ^{MOM}		ϕ^{VAL}	
$Spread_X$	High	Low	High	Low
0.05	0.314	0.337	0.381	0.410
0.1	0.304	0.349	0.368	0.425
0.2	0.284	0.378	0.344	0.463
0.3	0.267	0.415	0.323	0.508
0.4	0.253	0.455	0.303	0.557
0.5	0.239	0.512	0.285	0.622
	$N^{MOM}\phi^{MOM}$		$N^{VAL}\phi^{VAL}$	
$Spread_X$	High	Low	High	Low
0.05	0.329	0.320	0.400	0.390
0.1	0.335	0.314	0.405	0.382
0.2	0.341	0.303	0.413	0.370
0.3	0.347	0.291	0.420	0.355
0.4	0.354	0.273	0.425	0.334
0.5	0.359	0.256	0.428	0.311

[4] To ensure the robustness of our results, we also experimented with different values of z (6, 12) and j (3, 6, 12) in the extended model (results available upon request). Our model's key results remain unchanged. First, a larger amount of momentum capital (i.e., when X_t or Y_t is high) is associated with a larger return in period 0 (when arbitrageurs put their on trades), a smaller momentum profit in the short run, and a larger reversal in the longer run. Second, as we increase either $Spread_X$ or $Spread_Y$, the return spread between the high and low momentum-capital states becomes larger.

2.2 Momentum (Value) Returns with respect to $CoMOM$ ($CoVAL$)

To speak more directly to our empirical results, Appendix Table A1 Panel A compares periods with high arbitrage activity vs. periods with low arbitrage activity in our simulated data. Although we do not explicitly model the correlations among momentum or value stocks (given that we have a one-asset model), we can interpret $CoMOM$ and $CoVAL$ as measures of aggregate arbitrage demand for the momentum and value strategies at each point in time. As shown in Equation (3), arbitrage demand for the momentum strategy in our simple model is $N_t^{MOM}\phi_t^{MOM}abs(\Delta P_{t-1})$ (i.e., the number of momentum traders times their momentum trading intensity times the magnitude of the momentum signal), and that for value is $N_t^{VAL}\phi_t^{VAL}abs(\hat{P}_t - P_t)$ (i.e., the number of value traders times their value trading intensity times the magnitude of the value signal). We highlight that $CoMOM$ and $CoVAL$ reflect not only the aggregate trading intensity of momentum and value investors (as captured by $N_t^{MOM}\phi_t^{MOM}$ and $N_t^{VAL}\phi_t^{VAL}$, respectively), but also the magnitude of the momentum and value signals at that point in time. The correlation in our model between aggregate arbitrage demand and the momentum signal as well as that between aggregate arbitrage demand and the value signal is over 90%. In the data, $CoMOM$ is strongly associated with the formation period momentum spread (correlation of 0.16), and $CoVAL$ is strongly associated with the value spread (correlation of 0.17).

Consequently, we sort all simulated periods into two groups based on the absolute value of either $N_t^{MOM}\phi_t^{MOM}\Delta P_{t-1}$ or $N_t^{VAL}\phi_t^{VAL}(\hat{P}_t - P_t)$ and track momentum and value returns in the subsequent periods. If ΔP_{t-1} or $\hat{P}_t - P_t$ is negative, we multiple subsequent stock returns by -1 to reflect the fact that arbitrageurs have a short position in the stock in those situations. Moreover, to highlight the differences in return patterns associated with $CoMOM$ and $CoVAL$, we analyze momentum and value strategies one at a time.[5] This choice is driven by the fact that we have a one-asset model in which the momentum signal and value signal tend to be positively correlated. To see this, consider a positive fundamental shock to the stock. This results in a positive momentum signal, but since newswatchers underreact to the positive shock, this shock also results in a positive value signal.

The simulation results are shown in Appendix Table A1. Panel A presents momentum returns after periods of high and low momentum activity ($CoMOM$). We report momentum returns in period 0 (when momentum traders put on the trade), in period 1 (when momentum traders hold their positions), and periods 2-12 (the long-run). Across all γ_{MOM} values,

[5] In other words, we allow arbitrageurs to only condition on one of the two signals at a time.

momentum spreads in the formation period are larger in high $CoMOM$ periods than in low $CoMOM$ periods. There is also a larger reversal to the momentum strategy after periods with high $CoMOM$ than periods with low $CoMOM$, again across all γ_{MOM} values. Momentum returns in the holding period (j=1 in this case), interestingly, depend on the γ_{MOM} value. When momentum traders are relatively risk-averse (i.e., with a low γ_{MOM}), momentum returns in period 1 are larger in high $CoMOM$ periods than in low $CoMOM$ periods; when momentum traders are relatively risk-tolerant (i.e., with a high γ_{MOM}), the reverse is true.

Panel B of the same table presents value strategy returns after periods of high and low value activity ($CoVAL$). Again, we report value returns in period 0 (when value traders put on the trade), in period 1 (when they hold their positions), and periods 2-12 (the long-run). In both the holding and post-holding periods, value returns are strictly larger after high realizations of $CoVAL$ than low realizations of $CoVAL$. This is true for all γ_{VAL} values.

Overall, these findings are relatively intuitive. The idea that momentum can become an overreaction strategy (where long-horizon returns are negative) in our dynamic model can be foreshadowed by the comparative statics of Hong and Stein (1999). The reason is that as a positive-feedback strategy, momentum trading amplifies the signal: more momentum activity in period t, all else equal, makes the signal more attractive for period t+1. This naturally leads to an overreaction phenomenon. Of course, the attractiveness of our model is that we allow traders to dynamically condition the intensity of their momentum strategy on the observable component of the variation in the amount of momentum capital.

The notion that value strategies always have positive long-horizon returns in our model is also intuitive. To see this, note that arbitrageurs never "over-trade" the value signal (unlike in the momentum case) because their trading dampens the signal; in the extreme case, for instance, where arbitrageurs are risk neutral, their infinitely-aggressive trades completely eliminate the value signal and the expected return to the value strategy is reduced to zero (not negative!). Consequently, the equilibrium value signal is always a positive predictor of future returns, and so is value arbitrage activity, which is simply a linear (or any other increasing) function of the value signal.

3 Further Discussion of Appendix Tables

3.1 Time-Varying Momentum Return Skewness and Market Beta

Daniel and Moskowitz (2016) and Daniel, Jagannathan, and Kim (2019) study the non-normality of momentum returns with a particular focus on the negative skewness in momentum returns. Both papers argue that momentum crashes are forecastable.[6] Appendix Table A2 reports the extent to which comomentum forecasts time-series variation in the skewness of momentum returns. We examine both the skewness of daily returns (in months 1-3) and weekly returns (months 1-6 and months 1-12).

As shown in Panel A, the skewness of daily momentum returns is noticeably lower when comomentum is high. Indeed, the skewness of daily returns in the first three months of the holding period is monotonically decreasing in comomentum. The 20 percent of the sample that corresponds to low values of comomentum has subsequent momentum returns that exhibit daily return skewness of -0.126, while the 20 percent of the sample that corresponds to high values of comomentum has subsequent momentum returns with a skewness of -0.348. The difference is both economically and statistically significant. Panels B and C document that the negative skewness we find in long-short momentum strategies roughly comes from both sides of the trade. This finding is consistent with long-short momentum traders playing an important role in these markets.

Finally, in Panel D of the Table, we also examine the way the betas of momentum portfolios (as well as their long and short components) change in the year after portfolio formation. Consistent with previous research (Chen, Singal, and Whitelaw, 2016), we find that momentum portfolios tend to have betas that increase over the next year and that this increase is roughly attributable to both the long and short side of the trade. However, we find no evidence that this effect varies with comomentum as the related point estimate (-0.046) is statistically insignificant (t-statistic of -0.42). These conclusions continue to hold even after zeroing in on the loser side of the portfolio.

[6] Daniel and Moskowitz (2016) show that market declines and high market volatility forecast momentum crashes. Daniel, Jagannathan, and Kim (2018) estimate a hidden Markov model that helps identify those times where momentum strategies experience severe losses.

3.2 Institutional Ownership

If crowded trading is responsible for overreaction in momentum profits, then one expects that our findings should be stronger among those stocks that are more likely to be traded by arbitrageurs. Appendix Table A3 tests this idea by splitting stocks into two groups based on the level of institutional ownership (as of the beginning of the holding period). We can strongly reject the null hypothesis that the degree of predictability linked to comomentum is the same across the two subsamples. Specifically, the Years 1 and 2 predictability associated with moving from low to high $CoMOM$ subperiods among only low-institutional-ownership stocks is a statistically-insignificant -0.54% per month (t-statistic of -1.41). In stark contrast, moving from low to high $CoMOM$ states among only high-institutional-ownership stocks is associated with a predictable difference in Years 1 and 2 returns of -1.42% per month (t-statistic of -4.36). A test of the difference in these Years 1 and 2 return spreads across these two non-overlapping groups of stocks has a t-statistic of -3.91, which rejects the null at the 0.1% level of significance.

Given the results of Lee and Swaminathan (2000), we have also examined splitting the sample in a similar fashion based on turnover and the book-to-market ratio. In either case, we find no difference in comomentum's ability to forecast time-variation in momentum's long-run reversal.

3.3 Other Forecasting Variables

In Appendix Table A4, we document the inability of traditional variables to forecast momentum returns. Both formation-period spreads in the momentum characteristic (Panel A) and formation-period spreads in valuation ratios across winner and loser stocks (Panel B) do not predict abnormal holding period returns on momentum strategies.

3.4 Instrumented Comomentum

There are a number of possible drivers of time variation in momentum arbitrage activity. One possibility is that as the momentum strategy becomes more profitable (due to, for example, stronger underreaction on the part of noise traders), arbitrageurs optimally allocate more capital to the momentum trade. In that case, arbitrageurs can be viewed as a stabilizing force. This view might be categorized as a demand-side explanation. Another possibility is that additional capital is dedicated to the momentum strategy for reasons that are unrelated to the strategy's expected return—for instance, due to trading costs, funding conditions, and/or end investors' (and perhaps portfolio managers') misperception of expected momentum strategy returns—and thus arbitrageurs become a destabilizing force. This view might be categorized as a supply-side explanation.

Taken at face value, a demand-side explanation seems inconsistent with the facts. Comomentum does not forecast higher future momentum holding period returns. Instead, it forecasts lower holding period returns, higher volatility, and higher crash risk. Moreover, as a key message of the paper, comomentum forecasts more negative post-holding-period momentum returns, implying that more underreaction by noise traders is not the main driver of the increased activity.

Lou and Polk (2021) (henceforth LP) document in their Table II that time-series variation in comomentum may be linked to various direct but imperfect measures of the inputs going into the arbitrage process. One of these input variables is lagged momentum returns (in the year prior to the comomentum construction). LP show that more arbitrage activity flows to the momentum strategy when the strategy has recently performed well. However, there is no empirical evidence that the holding period returns on momentum strategies exhibit strong persistence. Nevertheless, some end investors and/or portfolio managers may hold this belief. This tension leads to a natural test of the supply side of arbitrage activity.

In particular, in Appendix Table A5, we repeat the analysis of LP's Table III but using a time series of fitted comomentum where the instrument is the cashflow-news component of lagged momentum returns.[7] We focus solely on this component of returns as doing so allows us to remove any feedback effect that momentum trading generates. Indeed, in a world where arbitrage activity in momentum stocks is just enough to eliminate underreaction but not so much as to result in overreaction, holding period returns to momentum strategies should consist of only cash-flow news.

The results from this test are consistent with our supply-side explanation and inconsistent with a demand-side explanation. As we show in Appendix Table A5, periods of high momentum returns are followed by a stronger reversal to the momentum strategy, consistent with a supply-side mechanism. In particular, both holding period and post-holding period returns to momentum stocks are markedly lower. Appendix Figure A8 shows these patterns in pre- and post-formation returns to momentum stocks using instrumented comomentum.

[7] We follow Vuolteenaho (2002) to extract the cashflow-news component of stock returns.

As the figure documents, we continue to find that arbitrage activity is destabilizing.

3.5 An Alternative Momentum Strategy

Since comomentum is a success at identifying times when arbitrage activity is high, we now examine whether our approach can help us identify arbitrage activity in the cross section. In particular, we develop trading strategies based on stocks' formation-year covariance with extreme momentum deciles. For every stock, we calculate the average of 1) its abnormal correlation with the returns of the top momentum decile and 2) the negative of its abnormal correlation with the bottom momentum decile in the formation year. We exclude, if necessary, that stock from the calculation of the decile returns. As with our comomentum measure, we industry adjust and control for the three factors of Fama and French (1993). We dub this measure stock comomentum ($CoMOM^{stock}$). We expect stock comomentum to identify those stocks that arbitrageurs are trading as part of their more general quantitative strategy.[8] These stocks should perform well subsequently and, if aggregate comomentum is high, eventually reverse.

Appendix Table A7 Panel A reports Fama-MacBeth estimates of cross-sectional regressions forecasting stock returns in month $t + 1$ with time $t - 1$ information (we skip the most recent month to avoid short-term return reversals). Regression (1) shows that stock comomentum strongly forecasts cross-sectional variation in monthly stock returns with a t-statistic over 4. We emphasize that stock comomentum is different from the typical measure of momentum risk sensitivity, i.e., the pre-formation loading on a momentum factor. To show this difference, we estimate the formation-period momentum beta ($beta\_UMD$) on Ken French's UMD factor using weekly returns over the same period in which we measure comomentum. Regression (2) shows that $beta\_UMD$ does not forecast cross-sectional variation in average returns. This failure is perhaps not surprising giving the literature emphasizing characteristics over covariances (Daniel and Titman, 1997).

Regression (3) documents that the momentum characteristic (MOM) works very well

[8] Wahal and Yavuz (2013) show that the past return on a stock's style predicts cross-sectional variation in average returns and that momentum is stronger among stocks that covary more with their style. Wahal and Yavuz measure style as the corresponding portfolio from the Fama and French (1993) 25 size and book-to-market portfolios. However, those portfolios may not be industry neutral (Cohen, Polk, and Vuolteenaho 2003) and may covary due to fundamentals (Cohen, Polk, and Vuolteenaho 2009). In contrast, our analysis not only focuses on comovement among momentum stocks (rather than stocks sorted on book-to-market and/or size) but also is careful to measure *excess* comovement, i.e., comovement controlling for the Fama and French (1993) market, size, and value factors.

over this time period. However, regression (4) shows that our stock comomentum measure remains significant in the presence of the momentum characteristic. Finally, regression (5) adds several other control variables including log size (ME), log book-to-market ratio (BM), idiosyncratic volatility ($IVOL$), and turnover ($TURNOVER$). Stock comomentum continues to be statistically significant.[9]

In Panel B of Appendix Table A7, we examine returns on a standard hedge portfolio based on stock-comomentum-sorted value-weight decile portfolios. Our goal with this simple approach is to confirm that the abnormal performance linked to stock comomentum is robust as well as to examine the buy-and-hold performance of the strategy. In particular, we report average (abnormal) monthly returns over months 1-6.[10] Results are economically and statistically significant in all instances. For example, the Fama and French three-factor alpha is 1% per month with an associated t-statistic of 4.07 when holding the long-short portfolio ranked by $CoMOM^{stock}$ for six months.

Finally, Panel C documents the ability of aggregate comomentum to forecast the returns on our stock comomentum strategy. As before, we classify all months into five groups based on $CoMOM$. In the row labeled "5-1", we report the difference in portfolio buy-and-hold returns over various horizons to the stock comomentum strategy based on investing in high comomentum periods (5) versus low periods (1). In the row labeled "OLS", we report the corresponding slope coefficient from the regression of the overlapping annual stock comomentum strategy returns (either in Year 0, 1, or 2) on comomentum ranks. Similar to what we find for the standard momentum strategy, the performance of the stock comomentum strategy is decreasing in aggregate comomentum, both in Year 1 and in Year 2.

3.6 International Evidence

As an out-of-sample test of our findings, we examine the predictive ability of comomentum in an international sample consisting of the returns to momentum strategies in the 19 largest markets (after the US). These countries are Australia (AUS), Austria (AUT), Belgium (BEL), Canada (CAN), Switzerland (CHE), Germany (DEU), Denmark (DNK),

[9] In unreported results, following Lee and Swaminathan (2000), we have also interacted $ret12$ with both $turnover$ and BM. These interactions have little effect on the ability of $CoMOM\_stock^L$ to describe cross-sectional variation in average returns.

[10] In unreported analysis, we find that the abnormal performance linked to stock comomentum lasts for six months. Then returns are essentially flat. Finally, all of the abnormal performance reverts in Year 2. These results are consistent with arbitrageurs causing overreaction that subsequently reverts.

Spain (ESP), Finland (FIN), France (FRA), Great Britain (GBR), Hong Kong (HKG), Italy (ITA), Japan (JPN), Netherland (NLD), Norway (NOR), New Zealand (NZL), Singapore (SGP), and Sweden (SWE). In each market, we calculate the country-specific comomentum measure in a manner similar to our US measure.

We find that our country-specific comomentum measures move together, with an average pairwise correlation of 0.47 over the subsample where we have data for all 19 countries (from December 1986 to December 2011). This finding is reassuring as one might expect that there is a common global factor in country-specific measures of arbitrage activity. Appendix Figure A9 plots equal-weight averages of the country-specific comomentum for each of three regions: Asia-Pacific, Europe, and North America. In the figure, North American comomentum declines very quickly after the 1987 crash and remains low until the late 1990s. The other two regions' comomentum declines slowly over this period. Then, all three regions' comomentum begins to move more closely together, generally increasing over the next 15 years. In general, there is a large common component in comomentum across countries and regions. For example, the correlation between comomentum in the US and Europe is 0.66, and the correlation between comomentum in the US and Japan is 0.19.

Appendix Table A8 Panel A reports the estimate from a regression forecasting a country's time t momentum return with time $t-1$ country-specific comomentum. Panel A also reports the regression coefficient after controlling for country-specific market, size, and value factors. We find that in every country these point estimates are negative. In particular, for the regression where we control for country-specific factors, six estimates have t-statistics less than -2.0, and 14 estimates have t-statistics less than -1.0. As a statistical test of the comomentum's forecasting ability in the international sample, we form a value-weight world momentum strategy (WLD) across these 19 non-US markets and forecast the resulting return with a corresponding value-weight comomentum measure, both without and with the corresponding global market, size, and value factors. The results confirm that comomentum strongly forecast future momentum returns in international settings, with t-statistics of -2.45 and -2.58 respectively.

If comomentum forecasts time-series variation in country-specific momentum and if our country comomentum measures are not perfectly correlated, a natural question to ask is whether there is cross-sectional (i.e., inter-country) information in our international comomentum measures. Thus, in Panel B of Appendix Table A8, we sort countries into quintiles based on their comomentum measure each month, investing in the momentum strategies of the countries in the bottom quintile and shorting the momentum strategies of the countries in the top quintile. We then adjust these monthly returns using world (including the US) market, size, value, and momentum factors.

We find that comomentum strongly forecasts the cross section of country-specific momentum strategy returns. Momentum strategies in low comomentum countries outperform momentum strategies in high comomentum countries by a factor of 3. (0.95% per month versus 0.29% per month) and the difference (0.66%) is statistically significant with a t-statistic of 2.98. These results continue to hold after controlling for market, size, and value factors. A strategy that only invests in momentum in those countries with low arbitrage activity and hedges out exposure to global market, size, and value factors earns more than 17% per year with a t-statistic of 8.41. Controlling for global momentum reduces this outperformance to a still quite impressive 4.2% per year which is statistically significant from zero (t-statistic of 2.61).

References

Chen Honghui, Singal, Vijay, and Robert F.Whitelaw, 2016, *Journal of Financial Economics* 121 624-644.

Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho, 2003, The Value Spread, *Journal of Finance* 58 609-641.

Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho, 2009, The Price is (Almost) Right, *Journal of Finance* 64 2739-2782.

Daniel, Kent and Tobias Moskowitz, 2016, Momentum Crashes, *Journal of Financial Economics* 122 221-247.

Daniel, Kent and Sheridan Titman, 1997, Evidence on the Characteristics of Cross Sectional Variation in Stock Returns, *Journal of Finance* 52, 1–33.

Daniel, Kent, Ravi Jagannathan, and Soohun Kim, 2019, Tail Risk in Momentum Strategy Returns, Columbia University working paper.

Fama, Eugene F. and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics* 33 3–56.

Grossman, Sanford, 1976, On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information, *Journal of Finance* 31, 573–585.

Hong, Harrison and Jeremy Stein, 1999, A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets, *Journal of Finance* 54 2143–2184.

Lee, Charles M. C. and Bhaskaran Swaminathan, 2000, Price Momentum and Trading Volume, *Journal of Finance* 55 2017–2069.

Lou, Dong and Christopher Polk, 2021, Comomentum: Inferring Arbitrage Activity from Return Correlations, *Review of Financial Studies*, Forthcoming.

Vuolteenaho, Tuomo, 2002, What Drives Firm-level Stock Returns?, *Journal of Finance* 57 233-264.

Wahal, Sunil and M. Denis Yavuz, 2013, Style Investing, Comovement, and Return Predictability, *Journal of Financial Economics* 107, 136-154.

Table A1: Model Comparative Statics

This table shows simulation results of our model. The model setup and the simulation procedure are discussed in the Online Appendix. Panel A shows returns to momentum traders after periods of low vs. high momentum activity (defined as $N_t^{MOM}\varphi_t^{MOM}abs(\Delta P_{t-1})$). Panel B shows returns to value traders after periods of low vs. high value activity (defined as $N_t^{VAL}\varphi_t^{VAL}abs(\hat{P}_t - P_t)$).

Panel A: Simulated Momentum Returns in the Model

	Low Momentum Activity			High Momentum Activity		
γ	Period 0	Period 1	Periods 2-12	Period 0	Period 1	Periods 2-12
1	0.1772	0.0458	-0.0570	0.5399	0.1100	-0.1567
3	0.1803	0.0309	-0.0706	0.5676	0.0566	-0.2264
5	0.1805	0.0244	-0.0767	0.5749	0.0363	-0.2441
7	0.1813	0.0221	-0.0792	0.5783	0.0268	-0.2482
10	0.1826	0.0201	-0.0820	0.5866	0.0188	-0.2647

Panel B: Simulated Value Returns in the Model

	Low Value Activity			High Value Activity		
γ	Period 0	Period 1	Periods 2-12	Period 0	Period 1	Periods 2-12
1	0.1929	0.0606	0.0544	0.7522	0.1760	0.2075
3	0.2142	0.0335	0.0457	0.8421	0.0931	0.1771
5	0.2225	0.0228	0.0403	0.8753	0.0636	0.1616
7	0.2269	0.0177	0.0399	0.8946	0.0472	0.1545
10	0.2309	0.0142	0.0395	0.9079	0.0326	0.1481

Table A2: Forecasting Momentum Return Skewness

This table reports the skewness of momentum returns as a function of lagged comomentum. At the end of each month, we sort all stocks into deciles based on their lagged 11-month cumulative returns (skipping the most recent month). We exclude stocks with prices below $5 a share and/or that are in the bottom NYSE size decile from the sample. We then classify all months into five groups based on $CoMOM$, the average of $CoMOM^L$ (loser comomentum) and $CoMOM^W$ (winner comomentum). $CoMOM^L$ is the pairwise abnormal return correlation in the loser decile in the ranking year t, while $CoMOM^W$ is the average pairwise abnormal return correlation in the winner decile. Panels A, B and C report the skewness in daily (weekly) returns to the value-weight winner minus loser portfolio in months 1 to 3 (1 to 6 and 1 to 12) after portfolio formation. Panel D reports the difference in market beta in the 12 months after portfolio formation vs. that during the formation year, based on weekly portfolio returns. Year zero is the portfolio ranking period. "5-1" is the difference in skewness of momentum returns (or changes in market beta) following high vs. low $CoMOM$. "OLS" is the slope coefficient from the regression of the skewness of momentum returns (or changes in market beta) on ranks of $CoMOM$. We compute t-statistics, shown in parentheses, based on standard errors corrected for serial-dependence with 3 lags (months 1-3), 6 lags (month 1-6), or 12 lags (month 1-12). We indicate statistical significance at the 5% level in bold.

Panel A: Momentum Skewness				
		Months 1-3	Months 1-6	Months 1-12
Rank	No Obs.	Estimate	Estimate	Estimate
1	112	-0.126	-0.209	-0.234
2	113	-0.140	-0.249	-0.258
3	113	-0.190	-0.343	-0.180
4	113	-0.250	-0.316	-0.271
5	113	-0.348	-0.423	-0.444
5-1		**-0.221**	-0.214	-0.210
		(-2.54)	(-1.74)	(-1.19)
OLS		**-0.055**	-0.050	-0.043
		(-2.94)	(-1.82)	(-1.16)

Panel B: Loser Skewness				
		Months 1-3	Months 1-6	Months 1-12
Rank	No Obs.	Estimate	Estimate	Estimate
1	112	0.245	0.244	0.263
2	113	0.185	0.273	0.269
3	113	0.259	0.364	0.247
4	113	0.302	0.321	0.335
5	113	0.366	0.636	0.685
5-1		0.121	**0.392**	0.422
		(1.24)	(2.83)	(1.84)
OLS		0.036	**0.083**	0.091
		(1.64)	(2.73)	(1.87)

Panel C: Winner Skewness				
		Months 1-3	Months 1-6	Months 1-12
Rank	No Obs.	Estimate	Estimate	Estimate
1	112	-0.012	-0.053	-0.045
2	113	-0.044	-0.124	-0.097
3	113	-0.103	-0.219	-0.099
4	113	-0.133	-0.216	-0.161
5	113	-0.228	-0.103	-0.087
5-1		**-0.216**	-0.050	-0.043
		(-2.46)	(-0.47)	(-0.41)
OLS		**-0.052**	-0.019	-0.015
		(-2.73)	(-0.79)	(-0.65)

Panel D: Market Beta Change				
		WML	Winners	Losers
Rank	No Obs.	Estimate	Estimate	Estimate
1	112	0.234	0.169	-0.065
2	113	0.217	0.127	-0.089
3	113	0.223	0.132	-0.092
4	113	0.178	0.103	-0.075
5	113	0.188	0.077	-0.111
5-1		-0.046	-0.092	-0.046
		(-0.42)	(-1.11)	(-0.71)
OLS		-0.013	-0.021	-0.008
		(-0.55)	(-1.18)	(-0.54)

Table A3: Institutional Ownership and the Comomentum Effect

This table reports the way the relation between expected returns on momentum stocks and lagged comomentum varies with institutional ownership. At the end of each month, we sort all stocks into deciles based on their lagged 11-month cumulative returns (skipping the most recent month). We exclude stocks with prices below $5 a share and/or that are in the bottom NYSE size decile from the sample. We then classify all months into five groups based on $CoMOM$, the average of $CoMOM^L$ (loser comomentum) and $CoMOM^W$ (winner comomentum). $CoMOM^L$ is the pairwise abnormal return correlation in the loser decile in the ranking year t, while $CoMOM^W$ is the average pairwise abnormal return correlation in the winner decile. We report the returns to the momentum strategy (i.e., long the value-weight winner decile and short the value-weight loser decile) in each of the four years after portfolio formation post-1980, following low to high $CoMOM$. Year zero is the portfolio ranking period. "5-1" is the difference in monthly returns to the momentum strategy following high vs. low $CoMOM$. "OLS" is the slope coefficient from the regression of monthly momentum returns on ranks of $CoMOM$. Panels A and B report the average monthly returns to the momentum strategy constructed solely based on stocks with low and high institutional ownership (banks and investment managers, as of the beginning of the holding period), respectively. Panel C compares the difference between "5-1" in the low institutional ownership subsample and "5-1" in the high institutional ownership subsample. We compute t-statistics, shown in parentheses, based on standard errors corrected for serial-dependence up to 24 lags and indicate statistical significance at the 5% level in bold.

Panel A: Stocks with Low Institutional Ownership						
		Year 0	Year 1	Year 2	Years 1-2	Years 3-4
Rank	No Obs.	Estimate	Estimate	Estimate	Estimate	Estimate
1	74	9.82%	0.39%	0.07%	0.23%	-0.08%
2	74	9.95%	0.78%	-0.27%	0.26%	-0.06%
3	74	11.08%	0.58%	-1.02%	-0.22%	0.24%
4	74	11.73%	0.66%	-0.77%	-0.05%	0.08%
5	74	12.47%	-0.52%	-0.11%	-0.31%	0.22%
5-1		2.65%	-0.91%	-0.18%	-0.54%	0.30%
		(1.91)	(-1.76)	(-0.45)	(-1.41)	(1.18)
OLS		**0.007**	-0.002	-0.001	-0.001	0.001
		(2.25)	(-1.48)	(-0.82)	(-1.76)	(1.08)

Panel B: Stocks with High Institutional Ownership						
		Year 0	Year 1	Year 2	Years 1-2	Years 3-4
Rank	No Obs.	Estimate	Estimate	Estimate	Estimate	Estimate
1	74	8.79%	0.60%	0.46%	0.53%	0.09%
2	74	9.10%	0.70%	-0.24%	0.23%	0.08%
3	74	10.06%	0.34%	-0.44%	-0.05%	-0.05%
4	74	10.61%	0.10%	-0.92%	-0.41%	-0.21%
5	74	11.66%	-0.70%	-1.08%	-0.89%	0.33%
5-1		**2.87%**	**-1.30%**	**-1.54%**	**-1.42%**	0.24%
		(2.92)	(-2.10)	(-5.63)	(-4.36)	(0.84)
OLS		**0.007**	**-0.003**	**-0.004**	**-0.003**	0.000
		(3.09)	(-2.38)	(-3.89)	(-4.55)	(0.26)

Panel C: Diff between High IO 5-1 and Low IO 5-1						
	No Obs.	Year 0	Year 1	Year 2	Years 1-2	Years 3-4
Diff	74	0.22%	-0.39%	**-1.36%**	**-0.88%**	-0.06%
		(0.51)	(-1.36)	(-3.22)	(-3.91)	(-0.15)

Table A4: Inability of Traditional Variables to Forecast Momentum Returns

This table reports momentum returns as a function of the lagged momentum formation period spread or momentum valuation spread. At the end of each month, we sort all stocks into deciles based on their lagged 11-month cumulative returns (skipping the most recent month). We exclude stocks with prices below $5 a share and/or that are in the bottom NYSE size decile from the sample. We then classify all months into five groups based on either the momentum formation period spread (i.e., the return spread between the winner and loser deciles during the formation period) or the momentum valuation spread (i.e., the difference in the average book-to-market ratio of the winner and loser deciles at the end of the formation period). We report the returns to the momentum strategy (i.e., long the value-weight winner decile and short the value-weight loser decile) in the month after portfolio formation. Year zero is the portfolio ranking period. "5-1" is the difference in monthly returns to the momentum strategy following high vs. low momentum formation-period spread or momentum valuation spread. "OLS" is the slope coefficient from the regression of monthly momentum returns on ranks of the two sorting variables. We compute t-statistics, shown in parentheses, based on standard errors corrected for serial-dependence with 12 lags and indicate statistical significance at the 5% level in bold.

Panel A: Sort by the Momentum Formation-Period Spread

		Month 1 Raw	Month 1 FF3F	Month 1 FF5F
Rank	No Obs.	Estimate	Estimate	Estimate
1	112	1.52%	1.87%	1.51%
2	113	0.65%	0.93%	0.83%
3	113	2.22%	2.30%	2.02%
4	113	0.66%	1.09%	1.13%
5	113	1.86%	2.25%	1.67%
5-1		0.34%	0.38%	0.15%
		(0.33)	(0.40)	(0.17)
OLS		0.001	0.001	0.001
		(0.29)	(0.41)	(0.30)

Panel B: Sort by the Momentum Valuation Spread

		Month 1 Raw	Month 1 FF3F	Month 1 FF5F
Rank	No Obs.	Estimate	Estimate	Estimate
1	112	0.01%	0.41%	0.21%
2	113	2.05%	2.11%	1.82%
3	113	2.12%	2.35%	1.91%
4	113	1.56%	2.17%	2.12%
5	113	1.16%	1.39%	1.08%
5-1		1.15%	0.99%	0.87%
		(1.63)	(1.37)	(1.19)
OLS		0.002	0.002	0.002
		(1.03)	(1.12)	(1.12)

Table A5: Instrumented Comomentum

This table shows returns to the momentum strategy as a function of an instrumented version of the comomentum measure that is based on the cashflow component of lagged momentum returns. At the end of each month, we sort all stocks into deciles based on their lagged 11-month cumulative returns (skipping the most recent month). We compute pairwise abnormal return correlations (after controlling for the Fama-French three factors) for all stocks in both the bottom and top deciles using weekly Fama-and-French 30-industry-adjusted stock returns in the previous 12 months. To mitigate the impact of microstructure issues, we exclude stock with prices below $5 a share and/or that are in the bottom NYSE size decile from the sample. We first calculate $CoMOM$, the average of $CoMOM^L$ and $CoMOM^W$, the average pairwise abnormal return correlation in the loser and winner deciles respectively in year t. We then classify all months into five groups based on the instrumented component of $CoMOM$. Reported below are the returns to a value-weight momentum strategy (i.e., winner minus loser deciles) in each of the four years after formation during 1965 to 2015, following low and high instrumented $CoMOM$. Year zero is the portfolio ranking period. "5-1" is the difference in monthly returns to the momentum strategy following high vs. low instrumented $CoMOM$. "OLS" is the slope coefficient from the regression of monthly momentum returns on ranks of instrumented $CoMOM$. We compute *t*-statistics, shown in parentheses, based on standard errors corrected for serial-dependence up to 24 lags and indicate statistical significance at the 5% level in bold.

Forecasting Future Momentum Returns						
		Year 0	Year 1	Year 2	Years 1-2	Years 3-4
Rank	No Obs.	Estimate	Estimate	Estimate	Estimate	Estimate
1	112	9.08%	0.95%	0.05%	0.50%	0.44%
2	113	9.24%	0.44%	-0.25%	0.10%	0.14%
3	113	9.46%	0.72%	-0.41%	0.15%	-0.24%
4	113	9.78%	0.61%	-0.79%	-0.09%	-0.48%
5	113	10.23%	0.06%	-0.75%	-0.35%	-0.35%
5-1		1.15%	**-0.90%**	**-0.80%**	**-0.85%**	**-0.79%**
		(0.98)	(-2.31)	(-3.07)	(-3.41)	(-3.80)
OLS		0.003	-0.002	**-0.002**	**-0.002**	**-0.002**
		(1.14)	(-1.92)	(-3.15)	(-3.53)	(-4.28)

Table A6: Forecasting Value Returns and Skewness with $CoVAL$

This table reports return characteristics of the value strategy as a function of lagged $CoVAL$. At the end of each month, we sort all stocks into deciles based on their lagged book-to-market ratios. We exclude stocks with prices below $5 a share and/or that are in the bottom NYSE size decile from the sample. We then classify all months into five groups based on $CoVAL$, the average of $CoVAL^V$ (value covalue) and $CoVAL^G$ (growth covalue). $CoVAL^V$ is the pairwise abnormal return correlation in the value decile in the ranking year t, while $CoVAL^G$ is the average pairwise abnormal return correlation in the growth decile. We report the returns to the value strategy (i.e., long the value-weight winner decile and short the value-weight loser decile) in each of the four years after portfolio formation during 1965 to 2015, following low to high $CoVAL$. Year zero is the portfolio ranking period. Panel A reports the average monthly return of the value strategy. Panel B reports the skewness in daily (weekly) returns to the value portfolio in months 1 to 3 (1 to 6 and 1 to 12) after portfolio formation. "5-1" is the difference in the relevant statistic across high and low $CoVAL$ ranks. "OLS" is the slope coefficient from the regression of the relevant statistic on ranks of $CoVAL$. We compute t-statistics, shown in parentheses, based on standard errors corrected for serial-dependence up to 24 lags. We indicate statistical significance at the 5% level in bold.

Panel A: Value Returns				
		Year 0	Years 1-2	Years 3-4
Rank	No Obs.	Estimate	Estimate	Estimate
1	112	-1.51%	0.01%	0.51%
2	113	-2.08%	0.44%	0.28%
3	113	-2.25%	0.31%	0.22%
4	113	-2.33%	0.61%	0.52%
5	113	-3.93%	1.18%	0.42%
5-1		**-2.42%**	**1.17%**	-0.09%
		(-4.20)	(2.39)	(-0.34)
OLS		**-0.005**	**0.003**	0.000
		(-3.73)	(2.26)	(0.10)

Panel B: Value Skewness				
		Months 1-3	Months 1-6	Months 1-12
Rank	No Obs.	Estimate	Estimate	Estimate
1	112	0.029	-0.011	-0.014
2	113	0.186	0.146	0.153
3	113	0.074	0.166	0.076
4	113	0.137	0.086	0.030
5	113	0.170	0.151	0.091
5-1		0.141	0.162	0.105
		(1.33)	(1.12)	(0.55)
OLS		0.023	0.026	0.009
		(1.01)	(0.87)	(0.22)

Table A7: An Alternative Momentum Strategy

This table reports the return to trading strategies based on stocks' formation-year covariance with momentum stocks. Panel A reports Fama-MacBeth estimates of cross-sectional regressions forecasting stock returns in month t+1. At the end of each month t, we sort all stocks into deciles based on their past 11-month cumulative return (skipping the most recent month to avoid short-term return reversals and excluding micro-cap and low-price stocks to mitigate the impact of microstructure issues). The main independent variable is $CoMOM^{Stock}$, the average of (1) the abnormal return correlation between weekly returns of a stock and the *minus* weekly returns to the bottom momentum decile in the formation year, and (2) the abnormal return correlation between weekly returns of a stock and weekly returns to the top momentum decile in the formation year. We exclude, if necessary, the stock in question from the calculation of the decile returns. Other control variables include the formation-period momentum beta with regard to the weekly UMD factor ($BETA^{UMD}$), lagged one-year stock return (MOM), log size (ME), log book-to-market ratio (BM), idiosyncratic volatility ($IVOL$), and turnover ($TURNOVER$). Panel B reports the average monthly buy-and-hold returns over various horizons to a long-short $CoMOM^{Stock}$ strategy formed from monthly-rebalanced value-weight decile portfolios. Panel C documents the ability of $CoMOM$ to forecast the $CoMOM^{Stock}$ strategy. We classify all months into five groups based on $CoMOM$. "5-1" is the difference in portfolio buy-and-hold returns over various horizons to the $CoMOM^{Stock}$ strategy based on investing in high (5) vs. low (1) $CoMOM$ groups. "OLS" is the corresponding slope coefficient from the regression of $CoMOM^{Stock}$ returns on ranks of $CoMOM$. Standard errors in brackets are Newey-West adjusted with 12 lags (Year 0, Year 1, and Year 2) or 24 lags (Year 3-4). We denote significance at the 90%, 95%, and 99% level using *, **, ***.

Panel A: Fama-MacBeth Regressions					
DepVar = Stock Returns in Month t+1					
	[1]	[2]	[3]	[4]	[5]
$CoMOM^{Stock}_{t-1}$	0.026***			0.009**	0.009**
	[0.006]			[0.004]	[0.004]
$BETA^{UMD}_{t-1}$		0.002*		0.000	-0.000
		[0.001]		[0.000]	[0.001]
MOM_{t-1}			0.007***	0.006***	0.007***
			[0.001]	[0.001]	[0.001]
ME_{t-1}					-0.002***
					[0.000]
BM_{t-1}					0.002**
					[0.001]
$IVOL_{t-1}$					-0.005***
					[0.001]
$TURNOVER_{t-1}$					-0.001
					[0.001]
Adj-R^2	0.02	0.02	0.04	0.05	0.09
No. Obs.	223,158	223,158	223,158	223,158	223,158

Panel B: Portfolio Returns Ranked by $CoMOM^{Stock}$									
Decile	Excess Return	CAPM Alpha	FF Alpha	Excess Return	CAPM Alpha	FF Alpha	Excess Return	CAPM Alpha	FF Alpha
	Months 1			Months 1-3			Months 1-6		
10 - 1	**1.15%**	**1.20%**	**1.30%**	**0.99%**	**1.03%**	**1.16%**	**0.77%**	**0.80%**	**1.00%**
	(3.89)	(4.07)	(4.41)	(3.76)	(3.94)	(4.27)	(3.20)	(3.35)	(4.07)

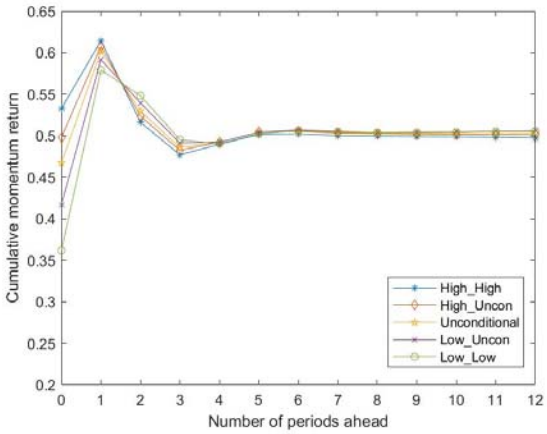
Panel C: Portfolio Returns Ranked by $CoMOM^{Stock}$ in Different *CoMOM* Periods								
	Year 0		Year 1		Year 2		Year 3-4	
Rank	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
5-1	**4.94%**	(4.26)	-1.05%	(-1.83)	**-1.27%**	(-2.64)	0.31%	(0.98)
OLS	**0.012**	(4.66)	**-0.003**	(-1.98)	**-0.003**	(-2.70)	0.000	(0.49)

Table A8: International Evidence

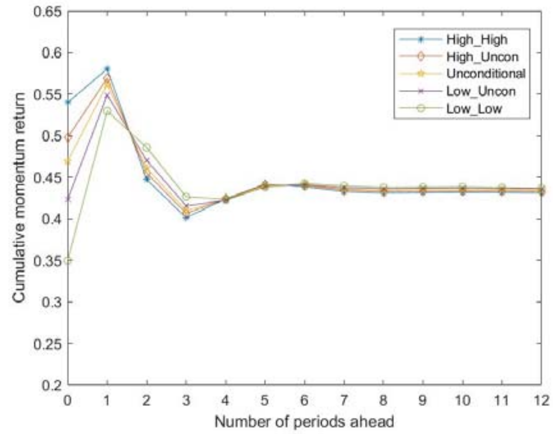
This table reports returns to international momentum strategies as a function of lagged country-specific comomentum. In Panel A, at the end of each month, we sort stocks in each market into deciles based on their lagged 11-month cumulative returns (skipping the most recent month). $CoMOM$ in each country is the average of $CoMOM^L$ (loser comomentum) and $CoMOM^W$ (winner comomentum). CoefEst1 is the regression coefficient of the month t momentum return on $CoMOM$ computed at the end of month t-1, while CoefEst2 is the corresponding regression coefficient, controlling for country-specific market, size, and value factors. We examine the world's largest 19 stock markets (after the US). We also compute a value-weight world (excluding the US) momentum strategy (WLD) and forecast that strategy with the corresponding value-weight world $CoMOM$ measure. In Panel B, we report the monthly returns to an inter-country (including the US) momentum timing strategy, which goes long (short) country-specific momentum strategies whose corresponding $CoMOM$ is in the bottom (top) quintile in the previous month. We then adjust these monthly returns using world (including the US) market, size, value, and momentum factors. We compute t-statistics, shown in parentheses, based on standard errors corrected for serial-dependence with 12 lags and indicate statistical significance at the 5% level in bold.

Panel A: Regression Coefficients in Other Countries							
Country	No months	CoefEst1	CoefEst2	Country	No months	CeofEst1	CeofEst2
AUS	324	**-0.0969**	-0.0634	GBR	324	-0.0585	-0.0536
		(-2.01)	(-1.37)			(-1.68)	(-1.78)
AUT	288	-0.051	-0.0627	HKG	324	**-0.0681**	**-0.1068**
		(-1.49)	(-1.02)			(-4.04)	(-2.38)
BEL	324	**-0.1043**	-0.0828	ITA	324	-0.0081	-0.0207
		(-2.18)	(-1.95)			(-0.30)	(-0.64)
CAN	324	**-0.1887**	**-0.1558**	JPN	324	-0.0354	-0.0427
		(-2.98)	(-3.16)			(-1.14)	(-1.69)
CHE	324	-0.0392	**-0.074**	NLD	324	**-0.0705**	**-0.0867**
		(-1.63)	(-2.41)			(-2.08)	(-2.04)
DEU	324	-0.0575	**-0.124**	NOR	321	-0.0172	-0.1075
		(-1.66)	(-2.46)			(-0.32)	(-1.77)
DNK	324	-0.0231	-0.0228	NZL	295	**-0.0837**	**-0.0643**
		(-0.98)	(-0.79)			(-2.45)	(-1.87)
ESP	324	-0.0003	0.0036	SGP	324	**-0.0899**	**-0.1428**
		(-0.01)	(0.10)			(-2.58)	(-4.21)
FIN	324	-0.0089	0.0054	SWE	324	-0.0065	0.0116
		(-0.26)	(0.17)			(-0.15)	(0.33)
FRA	324	-0.0711	-0.0628	WLD	324	**-0.0898**	**-0.0666**
		(-1.95)	(-1.86)			(-2.45)	(-2.58)

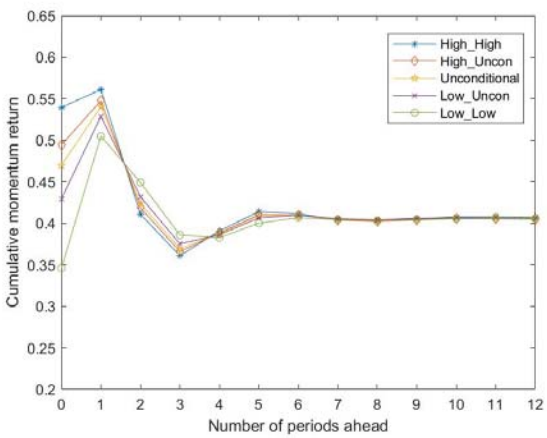
Panel B: Long-Short Portfolios of Country Momentum					
Quintile	No Months	Excess Return	CAPM Alpha	FF Alpha	Carhart Alpha
S	288	0.29%	0.41%	0.75%	-0.20%
		(1.11)	(1.73)	(3.49)	(-0.92)
L	288	0.95%	1.06%	1.45%	0.35%
		(3.90)	(4.65)	(8.41)	(2.61)
L-S	288	**0.66%**	**0.64%**	**0.71%**	**0.55%**
		(2.98)	(2.94)	(3.25)	(2.10)



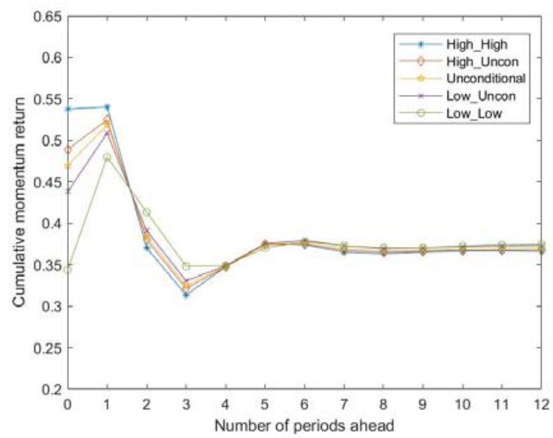
(a) $\gamma^{MOM} = 1$



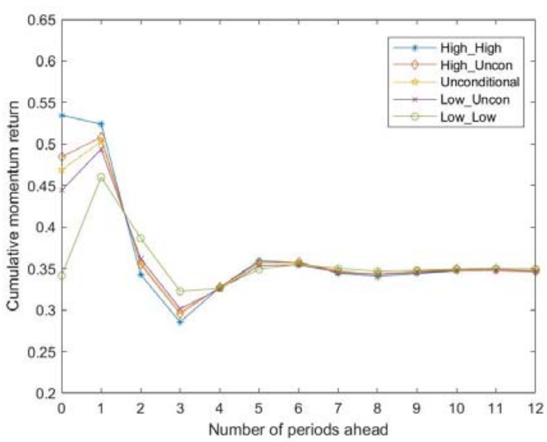
(b) $\gamma^{MOM} = 2$



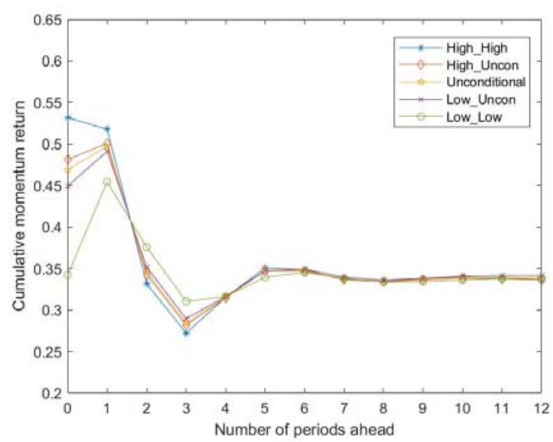
(c) $\gamma^{MOM} = 3$



(d) $\gamma^{MOM} = 5$

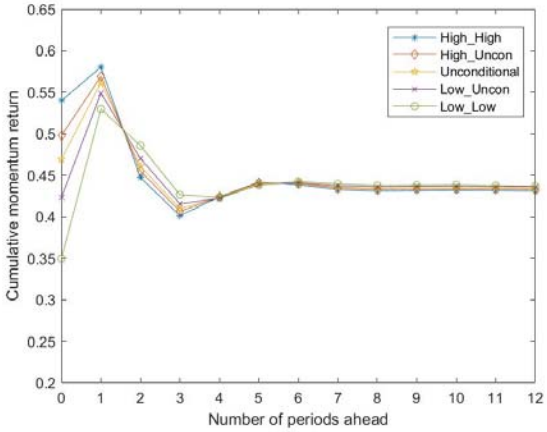


(e) $\gamma^{MOM} = 8$

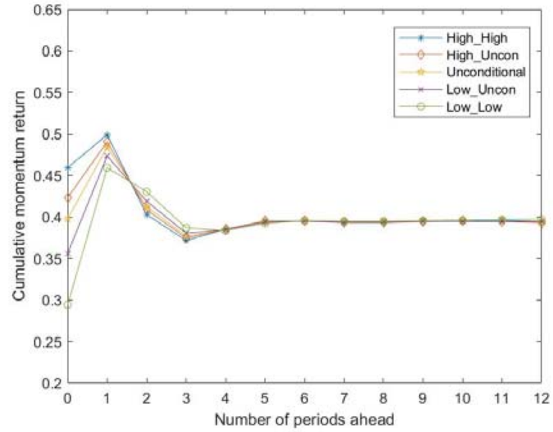


(f) $\gamma^{MOM} = 10$

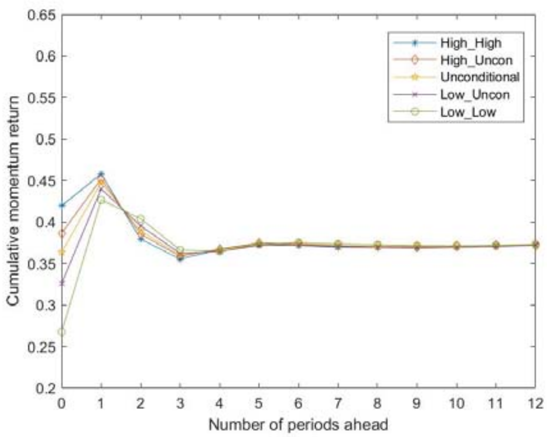
Figure A1: This figure shows the simulated regression coefficients of future stock returns on the momentum signal as a function of γ^{MOM} (the risk tolerance of momentum traders). The simulation is based on the baseline model (which has separate groups of arbitrageurs trading momentum and value). Arbitrageurs put on their trades in period 0.



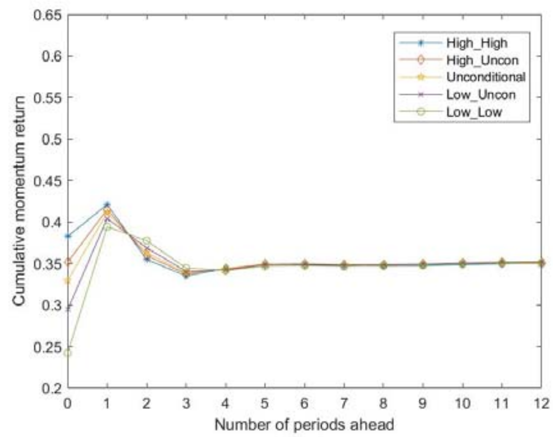
(a) $\gamma^{VAL} = 1$



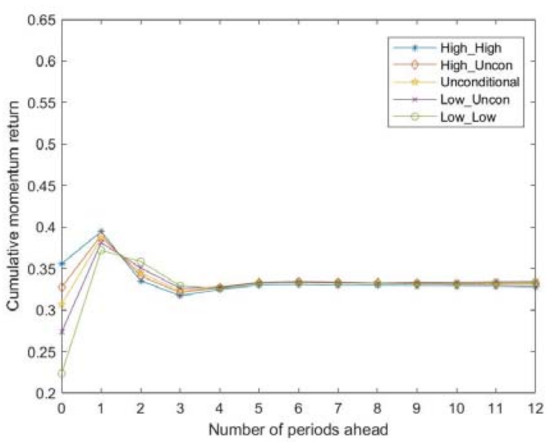
(b) $\gamma^{VAL} = 2$



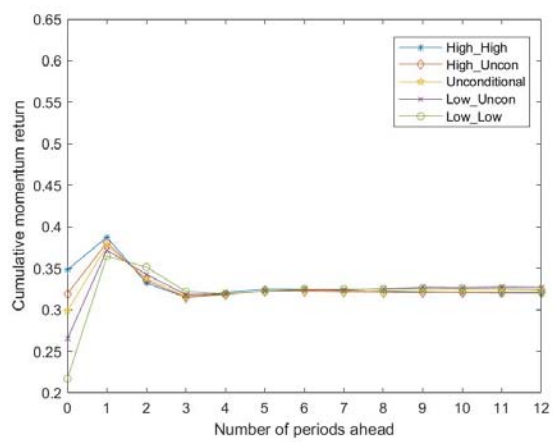
(c) $\gamma^{VAL} = 3$



(d) $\gamma^{VAL} = 5$

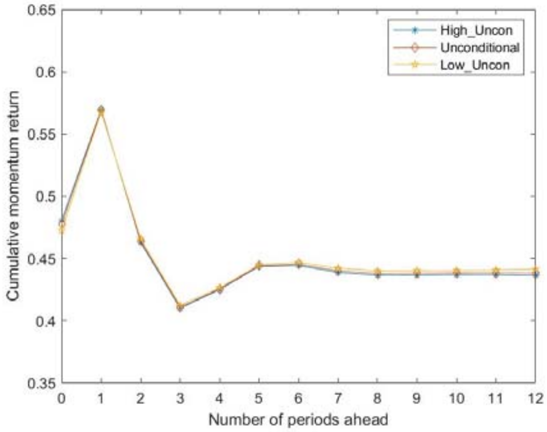


(e) $\gamma^{VAL} = 8$

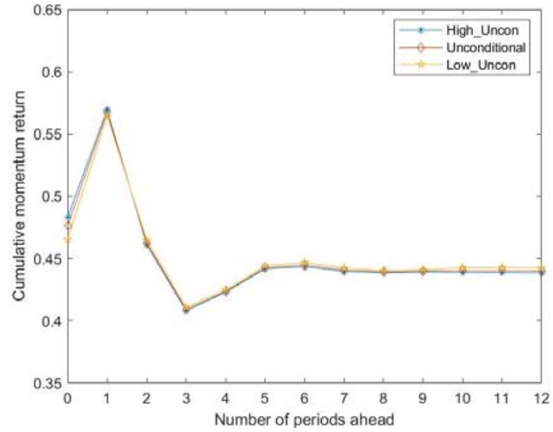


(f) $\gamma^{VAL} = 10$

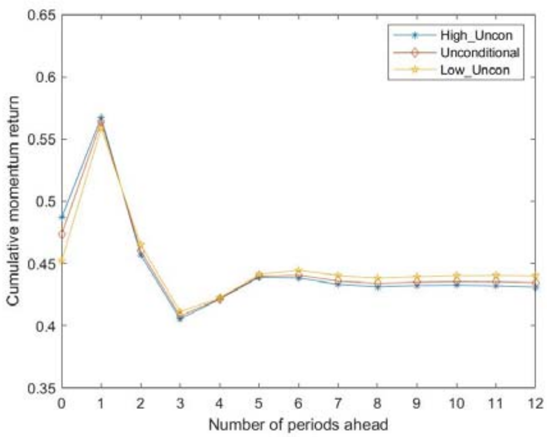
Figure A2: This figure shows the simulated regression coefficients of future stock returns on the momentum signal as a function of γ^{VAL} (the risk tolerance of value traders). The simulation is based on the baseline model (which has separate groups of arbitrageurs trading momentum and value). Arbitrageurs put on their trades in period 0.



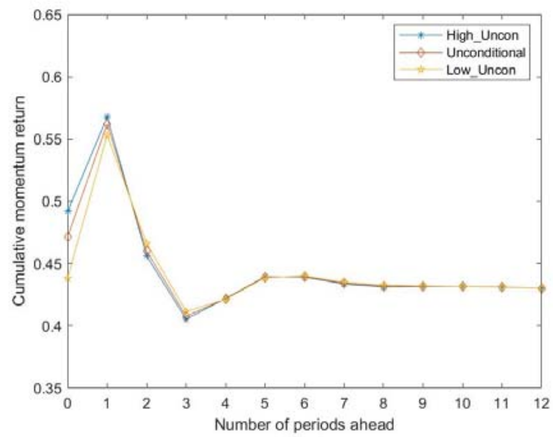
(a) $Spread_X = 0.05$



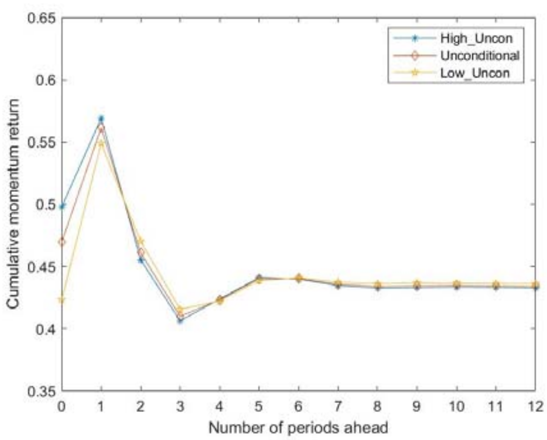
(b) $Spread_X = 0.1$



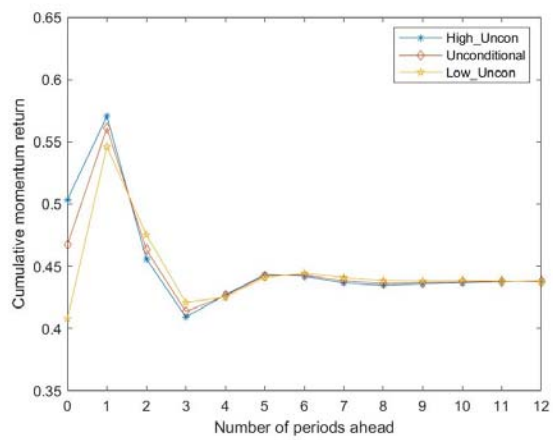
(c) $Spread_X = 0.2$



(d) $Spread_X = 0.3$

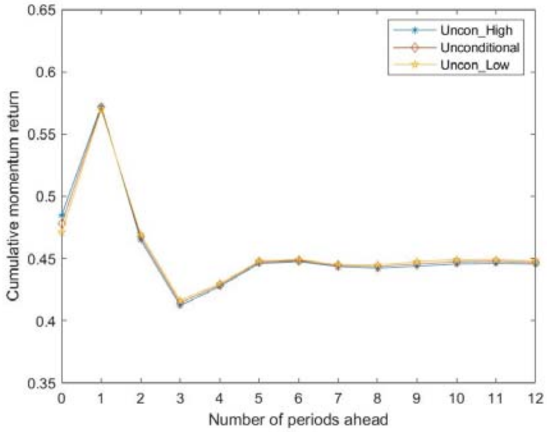


(e) $Spread_X = 0.4$

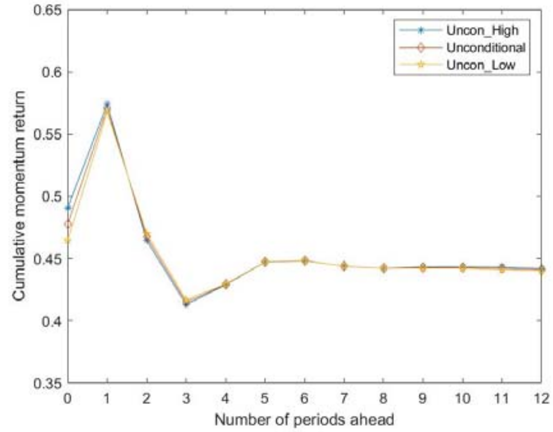


(f) $Spread_X = 0.5$

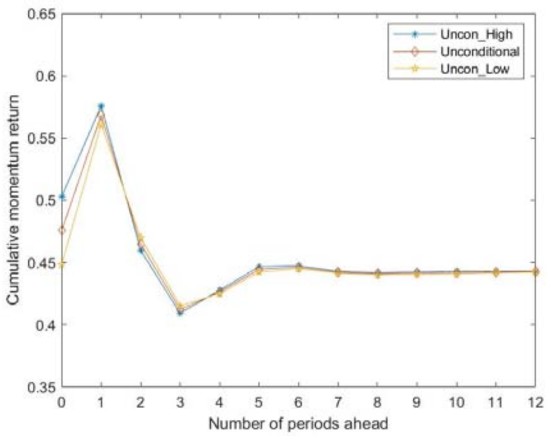
Figure A3: This figure shows the simulated regression coefficients of future stock returns on the momentum signal as a function of $Spread_X$ (the spread in the observable component of arbitrage capital). The simulation is based on the baseline model (which has separate groups of arbitrageurs trading momentum and value). Arbitrageurs put on their trades in period 0.



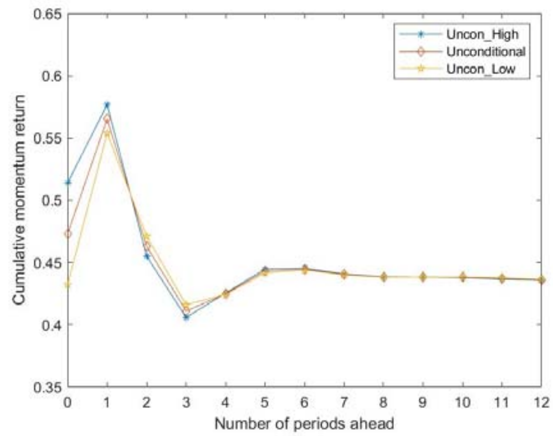
(a) $Spread_Y = 0.05$



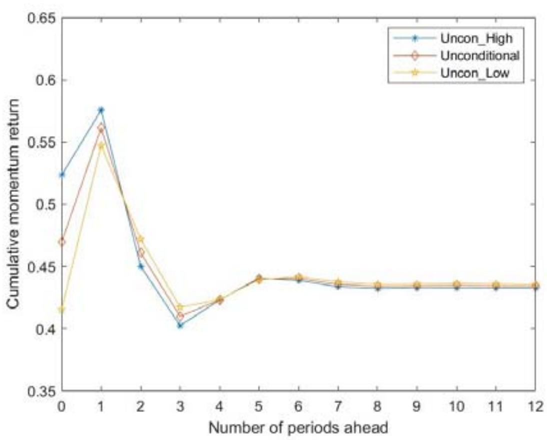
(b) $Spread_Y = 0.1$



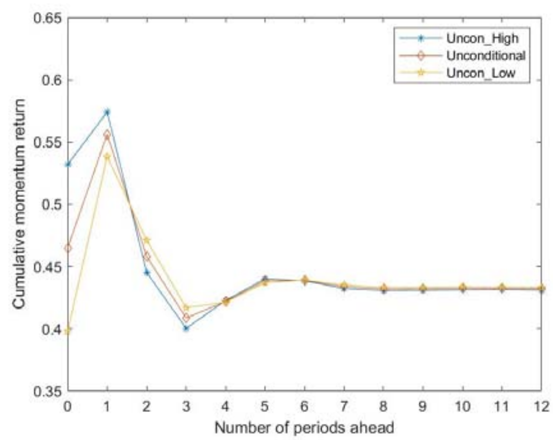
$Spread_Y = 0.2$



(d) $Spread_Y = 0.3$

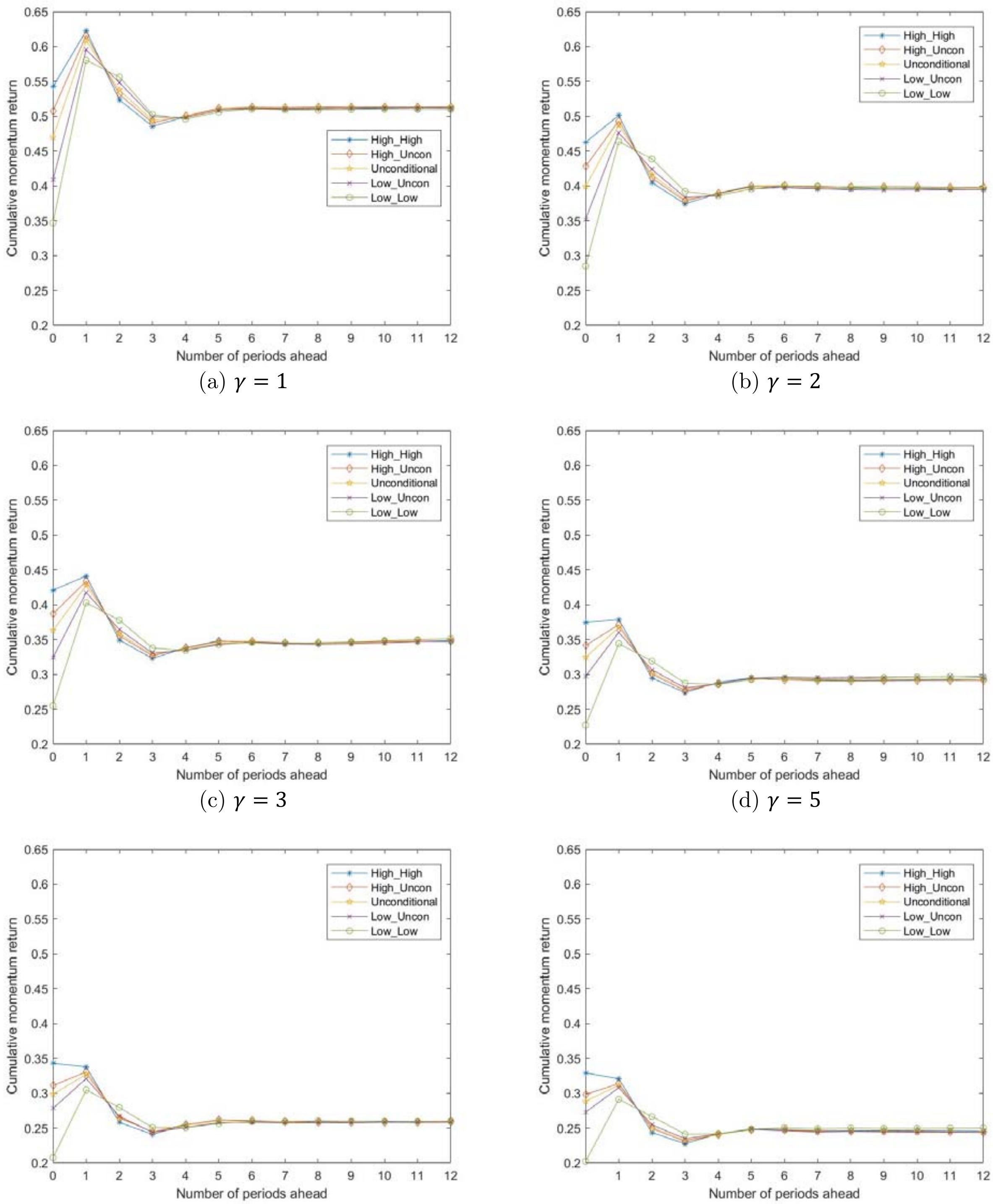


$Spread_Y = 0.4$



(f) $Spread_Y = 0.5$

Figure A4: This figure shows the simulated regression coefficients of future stock returns on the momentum signal as a function of $Spread_Y$ (the spread in the unobservable component of arbitrage capital). The simulation is based on the baseline model (which has separate groups of arbitrageurs trading momentum and value). Arbitrageurs put on their trades in period 0.

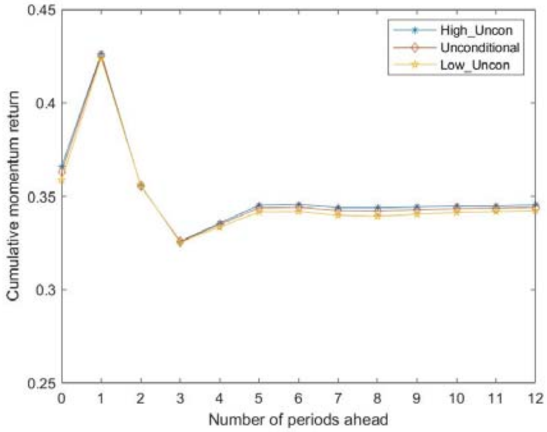


(a) $\gamma = 1$ (b) $\gamma = 2$

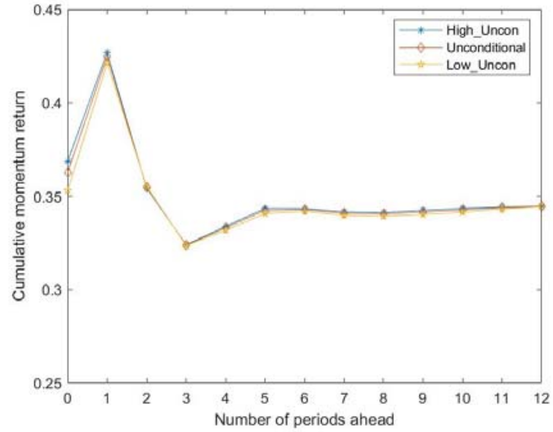
(c) $\gamma = 3$ (d) $\gamma = 5$

(e) $\gamma = 8$ (f) $\gamma = 10$

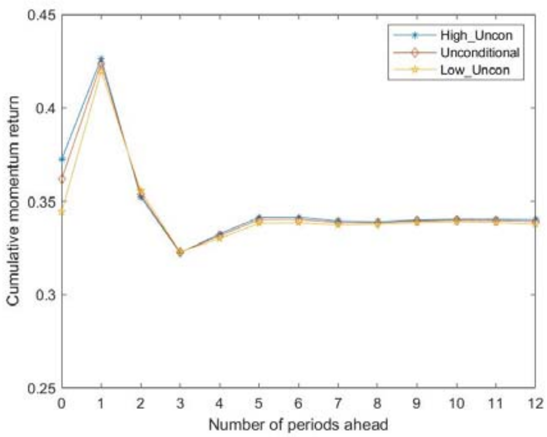
Figure A5: This figure shows the simulated regression coefficients of future stock returns on the momentum signal as a function of γ (the risk tolerance of all arbitrageurs). The simulation is based on the extended model (which has a single group of arbitrageurs that trade both the momentum and value signals optimally). Arbitrageurs put on their trades in period 0.



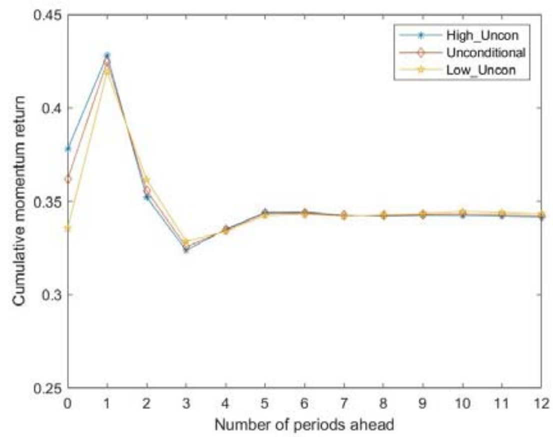
(a) $Spread_X = 0.05$



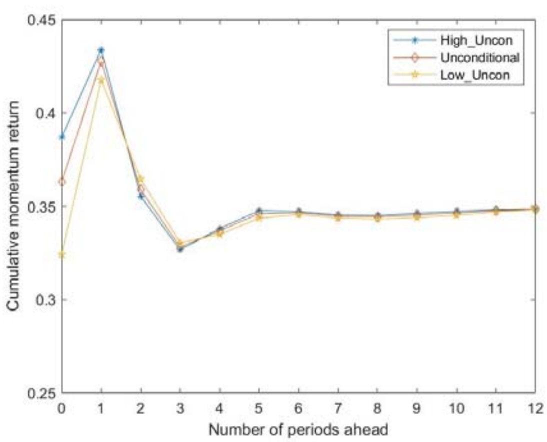
(b) $Spread_X = 0.1$



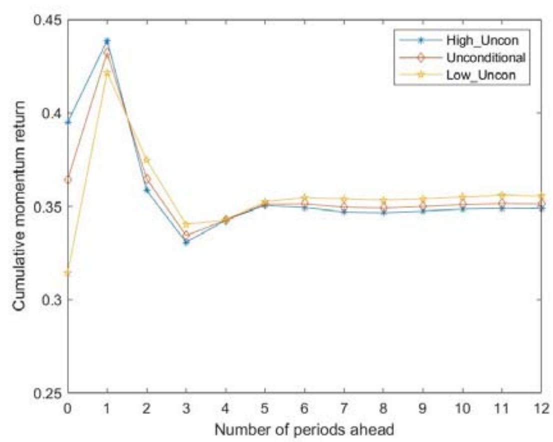
(c) $Spread_X = 0.2$



(d) $Spread_X = 0.3$

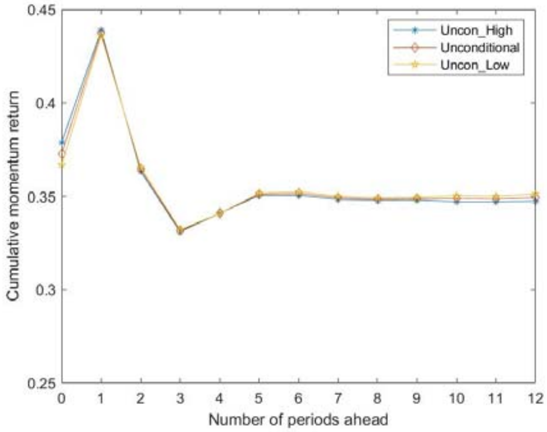


(e) $Spread_X = 0.4$

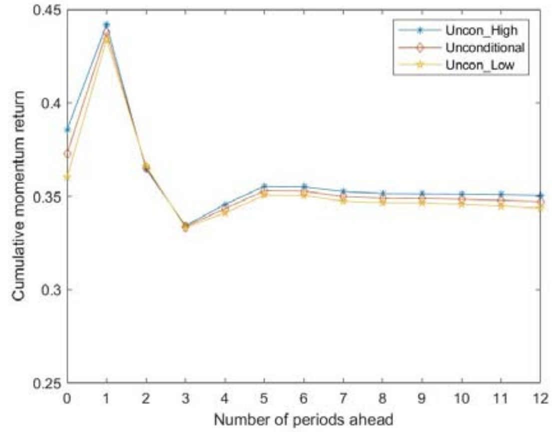


(f) $Spread_X = 0.5$

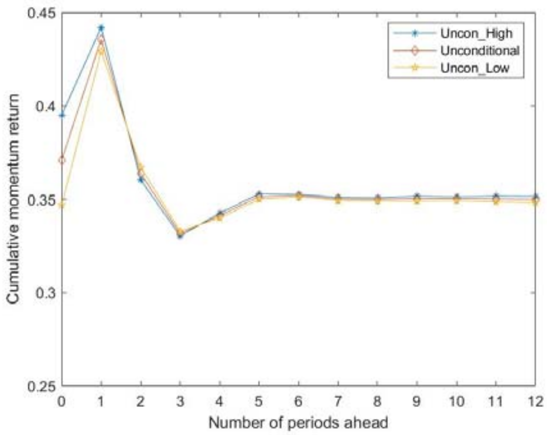
Figure A6: This figure shows the simulated regression coefficients of future stock returns on the momentum signal as a function of $Spread_X$ (the spread in the observable component of arbitrage capital). The simulation is based on the extended model (which has a single group of arbitrageurs that trade both the momentum and value signals optimally). Arbitrageurs put on their trades in period 0.



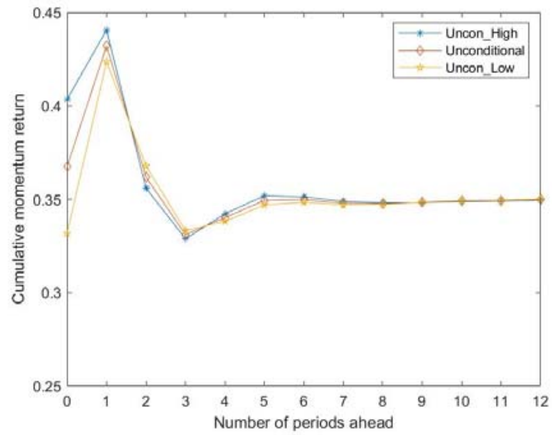
(a) $Spread_Y = 0.05$



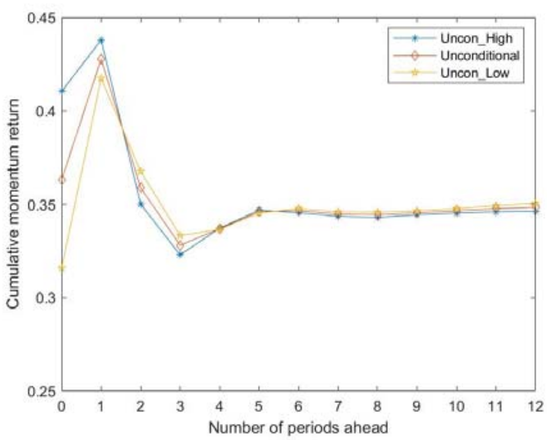
(b) $Spread_Y = 0.1$



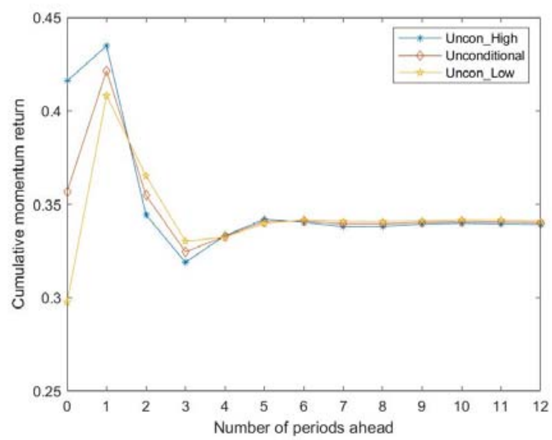
(c) $Spread_Y = 0.2$



(d) $Spread_Y = 0.3$

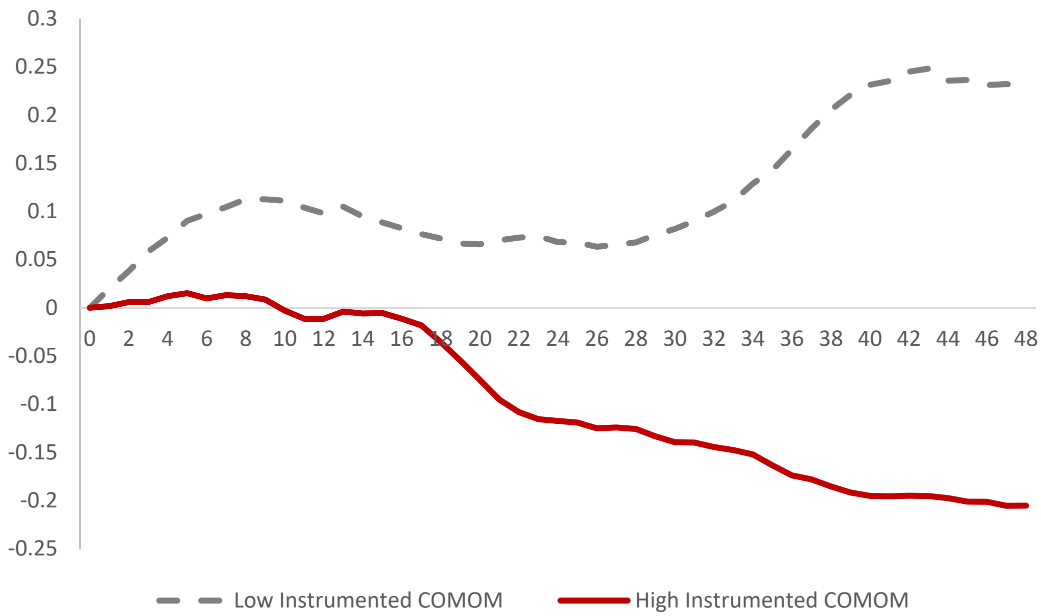


(e) $Spread_Y = 0.4$



(f) $Spread_Y = 0.5$

Figure A7: This figure shows the simulated regression coefficients of future stock returns on the momentum signal as a function of $Spread_Y$ (the spread in the unobservable component of arbitrage capital). The simulation is based on the extended model (which has a single group of arbitrageurs that trade both the momentum and value signals optimally). Arbitrageurs put on their trades in period 0.



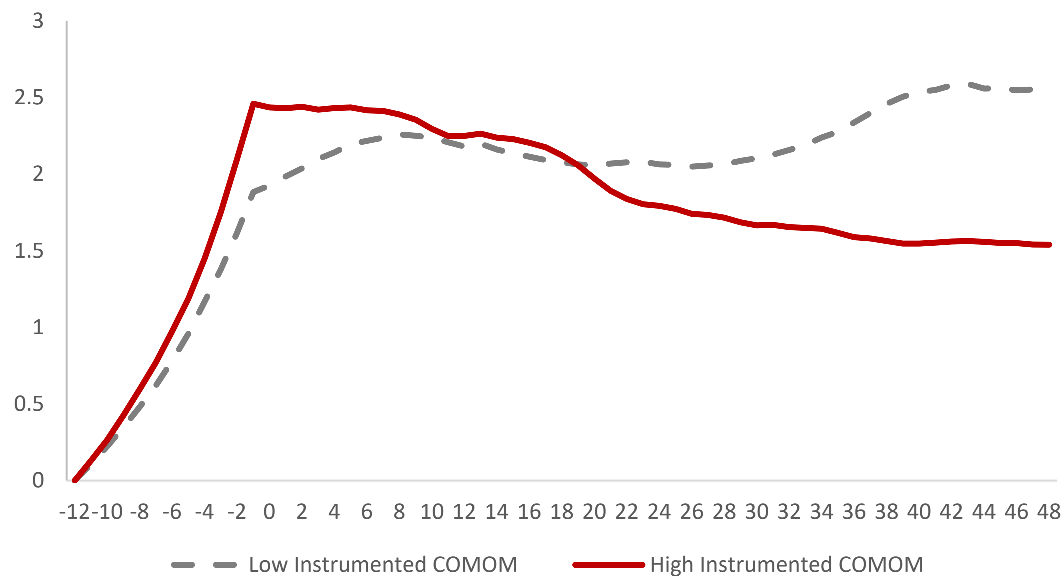


Figure A8: These figures show returns to the momentum strategy as a function of the instrumented comomentum measure (where the instrument is the cashflow component of lagged momentum holding-period returns). At the end of each month, we sort all stocks into deciles based on their lagged 11-month cumulative returns (skipping the most recent month). We compute pairwise abnormal return correlations (after controlling for the Fama-French three factors) for all stocks in both the bottom and top deciles using weekly Fama-and-French 30-industry-adjusted stock returns in the previous 12 months. To mitigate the impact of microstructure issues, we exclude stocks with prices below $5 a share and/or that are in the bottom NYSE size decile from the sample. We then classify all months into five groups based on the instrumented component of $CoMOM$, the average of $CoMOM^L$ and $CoMOM^W$. The top panel shows the compounded returns to a value-weight momentum strategy (i.e., winner minus loser deciles) in the four years after formation, following low and high instrumented $CoMOM$. The bottom panel shows the compounded returns to a value-weight momentum strategy (i.e., winner minus loser deciles) from the beginning of the formation year to four years post-formation following low and high instrumented $CoMOM$. The difference in years one and two abnormal returns following high vs. low $CoMOM$ periods is -0.85%/month with a t-statistic of -3.41.

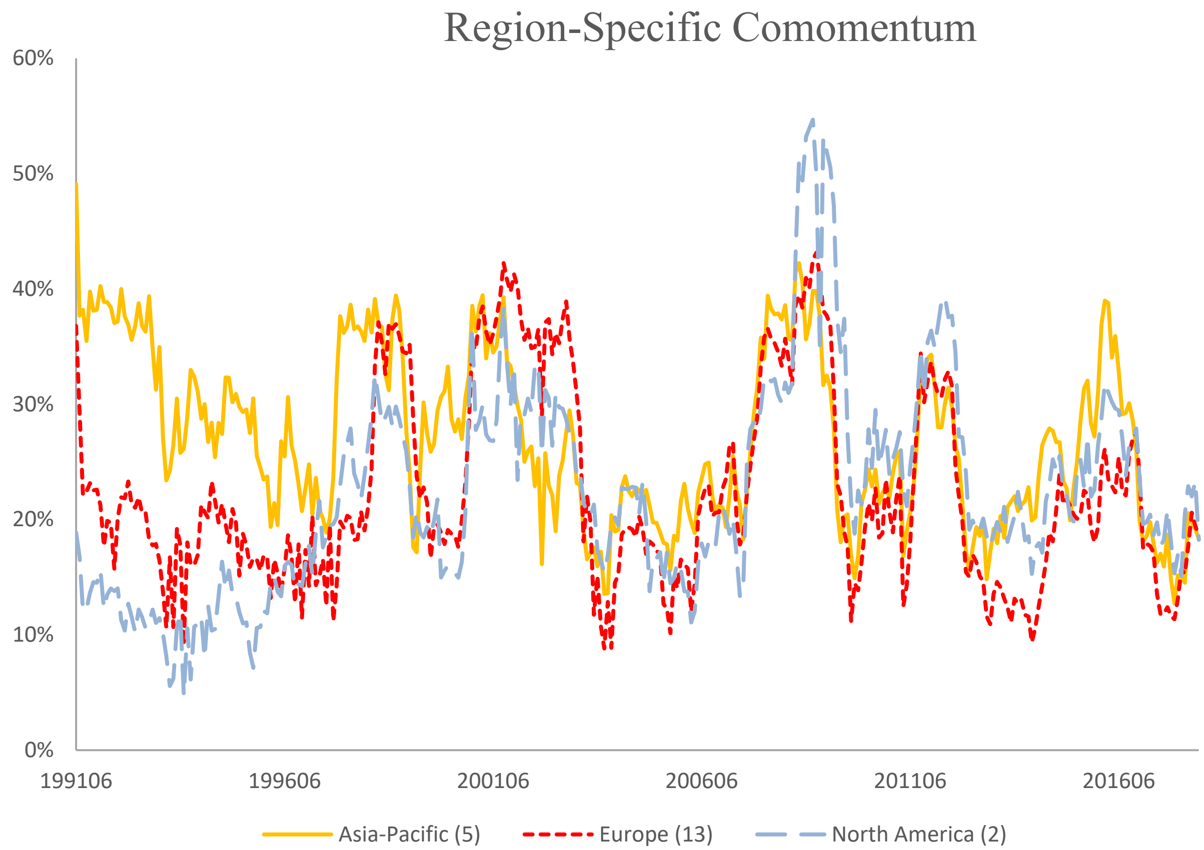


Figure A9: This figure shows the time series of region-specific comomentum measures. At the end of each month, we sort all stocks in a country into decile portfolios based on their lagged 11-month cumulative returns (skipping the most recent month). Country comomentum is the average pairwise return correlation in the loser and winner deciles measured in the ranking month. We calculate region comomentum as the equal-weight country momentum in the region. These regions are Asia-Pacific, Europe, and North America.