Internet Appendix

Fundamental Arbitrage under the Microscope: Evidence from Detailed Hedge Fund Transaction Data

Bastian von Beschwitz	Sandro Lunghi	Daniel Schmidt
Federal Reserve Board	Inalytics	HEC Paris

December 21, 2017

This internet appendix collects supplementary results and data descriptions for our paper. In Section A, we provide additional information on variable construction and on how we define investment regions. In Section B, we provide fund-level summary statistics and additional data descriptions. In Section C, we provide robustness checks to some regressions in the paper. In Section D, we show additional results that are referenced in the paper.

Section A: Additional Information on Dataset Construction

1) Regions

Following Karolyi and Wu (2014), we estimate DGTW returns and 4-factor alphas at the regional level as this provides for a reasonable compromise between a desirable granularity and the need to sufficiently populate 125 portfolios. As in Karolyi and Wu (2014), we categorize stock markets into 5 regions (Japan, North America, Europe, Asia-Pacific and Emerging Markets). All but the EME region are identical to Fama and French (2012). The assignment of countries into regions is displayed below in Table A.1.

Table A.1: Regions

Country Name	**Region**
Japan	Japan
Canada	North America
United States	North America
Australia	Asia-Pacific
New Zealand	Asia-Pacific
Singapore	Asia-Pacific
Hong Kong	Asia-Pacific
Austria	Europe
Belgium	Europe
Denmark	Europe
Finland	Europe
France	Europe
Germany	Europe
Greece	Europe
Ireland	Europe
Italy	Europe
Netherlands	Europe
Norway	Europe
Portugal	Europe
Spain	Europe
Sweden	Europe
Switzerland	Europe
United Kingdom	Europe
Argentina	EME
Brazil	EME
Chile	EME
China	EME
Colombia	EME
Czech Republic	EME
Hungary	EME
India	EME
Indonesia	EME
Israel	EME
Korea (South)	EME
Malaysia	EME
Mexico	EME
Pakistan	EME
Peru	EME
Philippines	EME
Poland	EME
Russian Federation	EME
South Africa	EME
Taiwan	EME
Thailand	EME
Turkey	EME
Venezuela	EME

2) Merging of datasets

We merge the trading and the holding datasets provided by Inalytics. We first merge based on ISIN. Trades that we cannot match by ISIN, we match by SEDOL and finally by CUSIP. Whenever there is a change in the number of shares held in the holdings data (and there was no stock split), we would expect to see a corresponding trade in the trade data. In fact, there are some errors in the data and the trade and holding data do not match perfectly. According to Inalytics, the holding data are more accurate. We therefore rely on the holdings data, i.e. we assume there is a trade whenever there is a change of holding in the holdings data. There are two exceptions to this: we adjust for some holdings that erroneously disappear and we make sure that stock splits (and stock dividends) are not identified as trades.

We treat as a mistake if a holding disappears from the data and then reappears shortly afterwards *without a trade being recorded*. In these cases we fill in the missing dates in between with the old holding quantity. Reappearing shortly afterwards means within 22 trading days (one month); or within 70 trading days (one quarter) if the position reappears with the exact same number of stocks. In total we identify 637 of these mistakes (compared to 150,000 trades in the full sample).

We identify stock splits in two ways: we use a dataset of corporate actions provided by Inalytics and we use Datastream data. Specifically, we assume that there is a stock split if shares outstanding in Datastream changed by at least 1% and there is a corresponding mismatch between the stock price change and the return (the Datastream return is adjusted for stock splits). We confirm the validity of the Datastream measure by confirming that it identifies over 95% of the stock splits from the corporate action data as stock splits. On days with a stock split we only treat holding changes as trades if they are initiating or closing trades (as these cannot come from a stock split). In total we identify 155 stock splits (compared to 150,000 trades in the full sample).

In total, we have about 150,000 (inferred) trades according to the holdings data. For about 90% of these trades, we have a corresponding trade in the trading data. However, for only about 83% of these trades does the number of stocks traded according to the trading data match the change in the number of stocks held in the holding data. In these cases we follow Inalytics' advice and assume the holdings data to be correct. In Table C.4 of this internet appendix, we show that our results are very similar if we only use those observations where the holdings and trade data perfectly agree.

3) Stock universe

To compute DGTW returns (and regional factors for the emerging market region, see below), we need a universe of stocks. We construct this stock universe by matching Worldscope and Datastream data. We only keep stocks that are covered in both databases. We only keep one stock per company (we identify companies using the Worldscope Permanent Identifier). We only keep stocks from the countries listed in Appendix A.1 (we take the country information from Worldscope). We require stocks to have a positive book value, information on market capitalization in Worldscope and a stock price of at least USD 0.20.

If funds trade stocks that are outside this stock universe (e.g., because they cannot be assigned to one of the regions or have no information on book value), we still include these trades in our sample. For such trades, we can only compute benchmark-adjusted returns and alphas (as described below) but we cannot compute DGTW returns. Our results are unchanged if we (1) exclude trades of stocks with a stock price of less than USD 1 (see Internet Appendix Table C.5) or (2) include only trades of stocks that are in the stock universe used to compute DGTW returns and factors (see Internet Appendix Table C.6).

4) Stock returns and balance sheet data

We download daily returns for stocks in our stock universe from Datastream using ISINs (and then using SEDOLs if we do not find a match using ISINs). We download returns in local currency and convert them to USD using the exchange rates on Datastream. Using local currency returns minimizes the errors due to rounding for stocks with low stock prices. When stocks are delisted, Datastream continues to report zero returns for these stocks. Following Busse, Goyal, Wahal (2014), we remove these trailing zeros, as well as any period with consecutives zero-return days that is at least 20 trading days long. When computing returns for the DGTW portfolios and Carhart (1997) factor, we remove returns in the top and bottom 0.25% by region following the instructions on the website of Kenneth French.

We take market capitalization in USD directly from Worldscope (code 07210) and compute book-to-market directly from Worlscope as the inverse of the price-to-book ratio (code 09304). We use annual Worldscope data.

For stocks in the Inalytics data that are not covered in Datastream, we receive stock return information from Inalytics. Because we don't have balance sheet information for these stocks, we cannot compute DGTW returns (but we can compute benchmark-adjusted returns and alphas). Of about 1.7 million stock-days in which a position is open, we have 1.43 million (84%) observations with return data on Datastream. By filling in the Inalytics return data we can increase this coverage to 1.66 million stock-days with returns (98%). We show in Table C.7 of this Internet Appendix that our results are robust to only using stock return data from Datastream.

5) Benchmark-adjusted returns

Benchmark-adjusted returns are defined as the stock return minus the return of the fund-specific benchmark index. The benchmark indexes are the benchmark returns against which hedge funds mark their own performance (for which they are then compensated). They are self-reported by the funds and do not change over the lifetime of a fund in our sample. Benchmark returns are provided to us by Inalytics.

6) Four-Factor alphas

To compute alphas, we use daily factor returns of the Carhart (1997) model for each of our 5 regions (see Table A.1 above). We use daily factors for America, Asia-Pacific, Europe, and Japan provided on Kenneth French's website. Because he does not provide factors for the emerging market region, we compute the emerging market factors ourselves following the instructions given on his website (our results are robust to excluding the EME region completely, see Table C.9 in this Internet Appendix). We use the U.S. 1-month T-bill rate as the risk free rate and all returns are in U.S. dollars. We compute market returns as value-weighted average returns for our stock universe in the EME region. To construct the SMB and HML factors, we sort stocks in the emerging market region into two market cap and three book-to-market equity (B/M) groups at the end of each June. Big stocks are those in the top 90% of (cumulative) market capitalization for the region, and small stocks are those in the bottom 10%. Fama and French (2012) use this method because for North America it roughly corresponds to the NYSE median used in Fama French (1993). According to Fama and French (2012), big stocks are more reliable for identifying B/M breakpoints. We follow their recommendation and set the B/M breakpoints for the emerging market region to the 30th and 70th percentiles of B/M for the big stocks in this region. For the 6 portfolios thus formed, we compute value-weighted returns for each day and then compute the factors as:

$$SMB = \frac{1}{3} * (Small\ Value + Small\ Neutral + Small\ Growth) - \frac{1}{3} * (Big\ Value + Big\ Neutral + Big\ Growth)$$

$$HML = \frac{1}{2} * (Small\ Value + Big\ Value) - \frac{1}{2} * (Small\ Growth + Big\ Growth)$$

The 2×3 sorts on size and lagged momentum to construct the MOM factor are formed monthly. For portfolios formed at the end of month t–1, the lagged momentum return is a stock's cumulative return for month t–12 to month t–2. The momentum breakpoints for the emerging market region are the 30th and 70th percentiles of the lagged momentum returns of the big stocks in the region. For the 6 portfolios thus formed, we compute value-weighted returns for each day and then compute the momentum factor as:

$$MOM = \frac{1}{2} * (Small\ High + Big\ High) - \frac{1}{2} * (Small\ Low + Big\ Low)$$

For each stock and each month, we then compute the beta with respect to their regional factors from a daily regression over the past year:

$$r_{c,t} - r_{f,t} = \alpha + \beta_m * (r_{m,t} - r_{f,t}) + \beta_{HML} * HML + \beta_{SMB} * SMB + \beta_{MOM} * MOM$$

where $r_{c,t}$ is the daily company return, $r_{m,t}$ is the daily market return and $r_{f,t}$ is the daily risk free rate.

For stocks that cannot be assigned to a region (either because country information is missing or the country is not included in any regions), we compute alphas relative to the global factors provided by Kenneth French. We show in Table C.9 of this Internet Appendix that our results are robust to excluding stocks that cannot be assigned to a region. We remove returns from the regression that are in the top and bottom 0.25% by region. Furthermore, we only keep betas that are based on at least 50 days of non-missing return data.

Following Frazzini and Pedersen (2014), we shrink the resulting beta estimates toward their cross sectional mean by computing:

$$\beta_{j,t}^{shrunk} = 0.7 * \beta_{j,t} + 0.3 * \bar{\beta}_{j,t}$$

for $j \in \{m, HML, SMB, MOM\}$ and where $\bar{\beta}_{j,t}$ is the equal-weighted average $\beta_{j,t}$ estimated in the region to which stock c belongs. Finally, we use shrunk betas to compute daily alphas as follows:

$$Four\ factor\ alpha_{c,t} = r_{c,t} - r_{f,t} - \beta_m * (r_{m,t} - r_{f,t}) - \beta_{HML} * HML_t - \beta_{SMB} * SMB_t - \beta_{MOM} * MOM_t$$

7) DGTW returns

To compute DGTW returns, we split the stocks in our universe into 625 portfolios. First, we split the universe into the 5 geographic region (see Table A.1 above). Second, each year, within each region we sort stocks into 5 portfolios by market capitalization. Third, within each of these 25 size-region portfolios we sort stocks by book-to-market. Fourth, within each of these 125 region-size-book to market portfolios, we sort stocks into 5 portfolios by returns over months t-12 to t-2. While splits for market cap and market-to-book happen once a year, splits by past return are executed every month.

We then compute the benchmark return for each of the 625 portfolio on each day as the value-weighted average return of all portfolio stocks (in USD). Finally, we compute DGTW returns as stock return minus the return of the respective benchmark portfolio.

8) Return winsorization

Since the international stock return data contains large outliers (Ince and Porter (2004)), we winsorize all our return measures at the 1% level on both sides.

Section B: Additional Data Description

1) Fund-specific information

The small size of the dataset allows us to provide information on each individual fund. At the same time, our data provider requires that we limit how much we disclose about each fund in order to preserve their anonymity. In Table B.1, we report the fund-level summary statistics that we can disclose. We report the number of short and long positions that a fund holds on average. The number of positions vary from 9 to 52 for short positions and from 13 to 211 for long positions. Funds also differ in how much they trade, from an average of 1 to 14 orders per day. Finally, we report the total number of orders and positions for each fund. There is not a single fund that dominates the dataset. Fund 16, which has the most orders and positions, accounts only for 15% of the orders and 13% of the positions in our sample.

Table B.1: Fund-specific information

This table displays summary statistics for each individual fund. *Number of Long (Short) Positions* is the average number of long (short) positions held per day. *Orders per Day* is the average number of orders executed per day. *Number of Orders in Dataset* is the total number of all orders of the fund in our dataset. *Number of Positions in Dataset* is the total number of different positions held by the fund. A position lasts from its opening—i.e., the first buy for long positions or the first sell for short positions—to its close—i.e., the moment when the stock holding goes back to zero.

Fund Number	Number of Long Positions	Number of Short Positions	Orders per Day	Number of Orders in Dataset	Number of Positions in Dataset
(1)	13	9	4.6	2,528	396
(2)	16	13	6.8	3,757	384
(3)	17	24	1	546	108
(4)	17	15	4.9	2,721	600
(5)	22	19	5.6	3,103	497
(6)	26	11	5.7	3,151	434
(7)	31	15	8.4	4,503	809
(8)	32	9	1.5	2,544	492
(9)	34	15	2.9	7,192	581
(10)	36	12	1.7	2,633	959
(11)	36	19	11.6	6,224	1,149
(12)	43	28	14	5,742	1,247
(13)	50	13	7.8	4,346	578
(14)	50	30	10.5	5,303	770
(15)	54	25	2.7	4,219	731
(16)	57	36	2.9	4,429	857
(17)	58	15	7	3,899	331
(18)	60	46	7.6	14,780	2,056
(19)	75	45	1	2,363	728
(20)	110	51	5	5,895	898
(21)	211	52	8.6	6,224	1,636

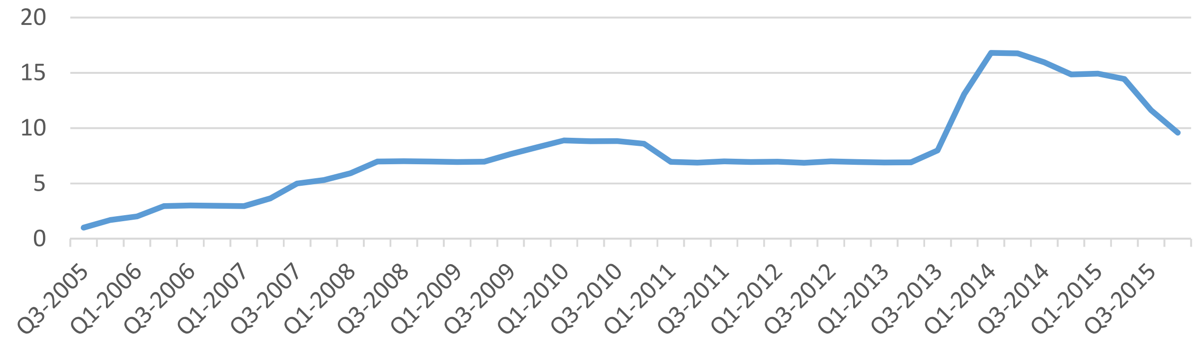
2) Coverage over Sample Period

Our sample period runs from 2005 to 2015. However, each individual fund covers only a fraction of this sample period. Figure B.2 below gives an overview about how the number of funds, positions and orders changes over the sample period. From 2005 to 2007, the sample is fairly small with only 1-6 funds. From late 2008 to mid-2013, we have 8 to 9 funds in the sample. In 2013, the number of funds jumps to 17. However, the early funds have more positions, so from 2008 Q1 to the end of the sample period we always have at least 500 open positions in the data. Orders move more proportional to the number of funds. From 2008 Q1 onward we have around 20 orders per day, but towards the end of the sample period that number jumps to over 100 orders per day. We include our full sample period in our tests to preserve statistical power and ensure that no specific time period is driving our results.

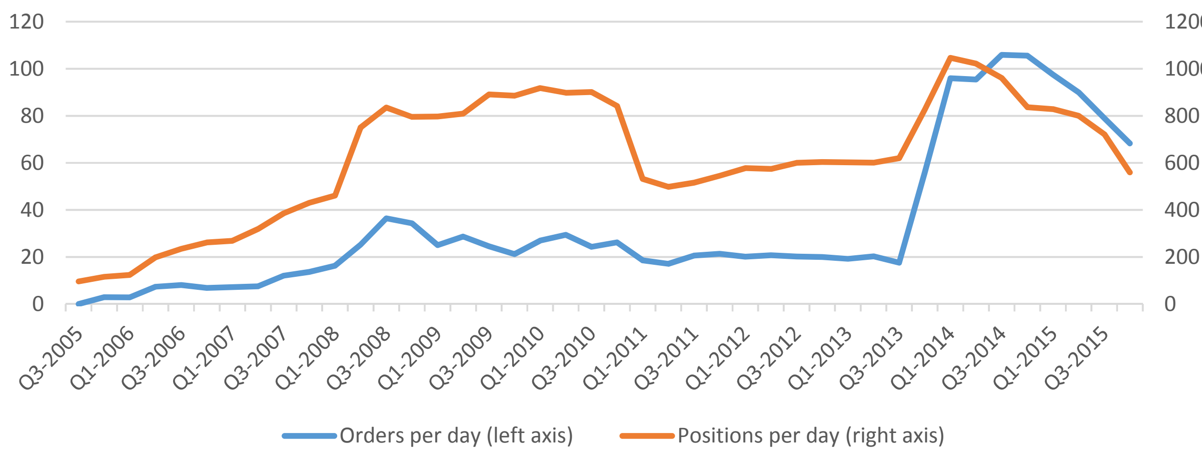
Figure B.2: Coverage over sample period

This figure shows the coverage over our sample period. Panel A shows the average number of funds in the sample for each quarter. Panel B shows the number of orders per day and of open positions per day averaged over the quarter.

Panel A: Number of funds in the sample



Panel B: Number of orders and positions per day



3) What stocks do our hedge funds invest in?

In this subsection, we study in more detail which type of stocks the hedge funds in our sample invest in. In In Figure B.3 Panel A, we plot the average fraction of shares outstanding held by our hedge funds for different market capitalization deciles. We see that this fraction is monotonically increasing with company size for long positions. For short positions, it is also generally increasing but peaks at the 9th decile. Also, short positions in small stocks (deciles 1 to 4) seem to be very rare, likely due to the difficulty of borrowing these stocks. In summary, similar to institutional investors in general (Gompers and Metrick (2001)), the hedge funds in our sample tend to focus on larger stocks.

Next, we examine whether our funds tend to concentrate their holdings in certain industries. In Figure B.3 Panel B, we plot the fraction of long and short positions that is held in each of the 12 Fama French industries. As a comparison, we also plot the fraction of total market capitalization concentrated in these industries. By and large, funds invest in all 12 industries in proportion to their market capitalization weights. If anything, funds tend to somewhat overweigh more traditional industries such as manufacturing, business equipment and retail, while they underweigh industries like finance and utilities that are subject to special rules and regulation.

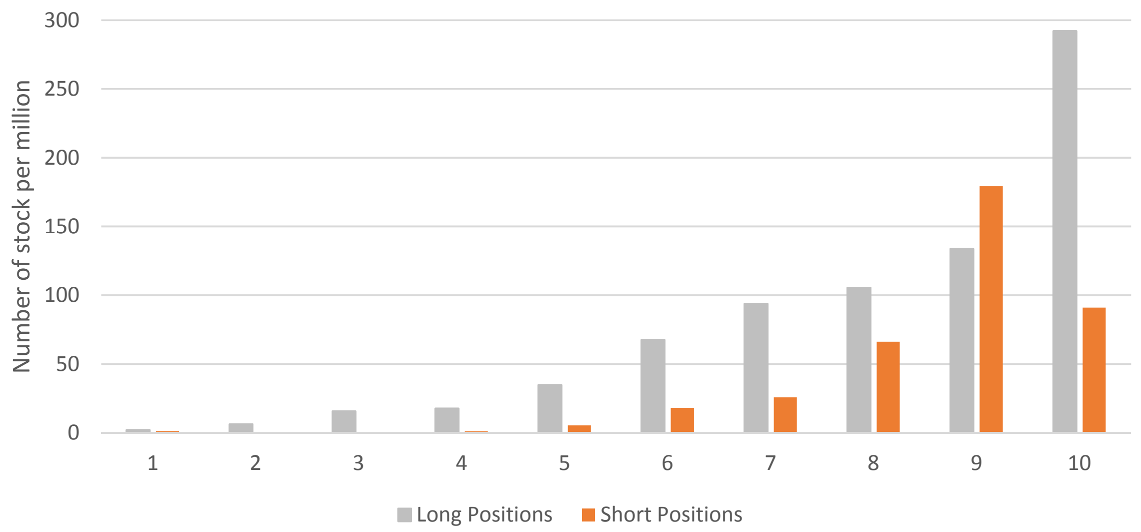
In Panel C, we present a similar plot for different deciles in terms of book-to-market ratio. Once again, the long and short position weights are fairly close to the market capitalization weights, although funds tend to somewhat overweigh growth stocks especially in long positions. Next, in Panel D, we present a plot for the different deciles of past 12-months return. Here, we observe a tendency of our funds to overweigh stocks with positive past returns in their long positions and to overweigh stocks with negative past returns in their short positions. This suggests that our funds engage in some momentum trading.

To conclude, we show that the long-short hedge funds in our sample spread their investments over many industries and different types of stocks. They tend to overweigh larger companies and engage in some momentum trading, but split fairly evenly across different industries and value vs. growth stocks.

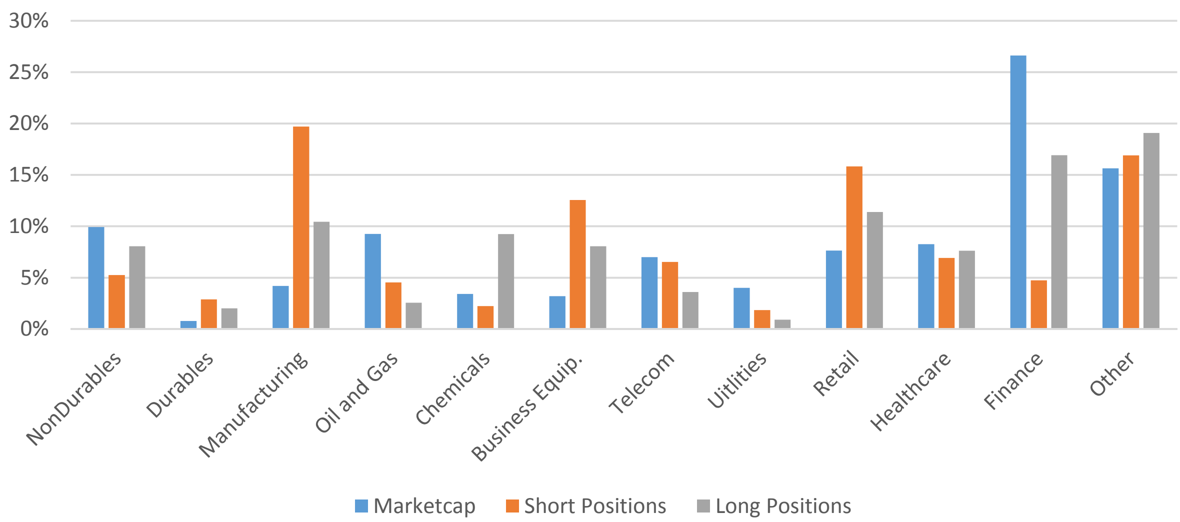
Figure B.3: Which stocks do the funds invest in?

This figure examines which stocks are more or less held by the hedge funds in our sample. In Panel A, we display the fraction of shares outstanding held by our hedge funds as long or short positions across different deciles of stocks in terms of market capitalization. The fraction is displayed in number of shares held per million of shares outstanding. We compute this fraction for each stock-month observation and then compute averages of this fraction. In Panel B, we display stock holdings by industry (using the 12 Fama French industry classification). Here, we display holding fractions separately for long and short positions. For comparison, we include the fraction of total market capitalization concentrated in each industry. We base the fraction of long and short positions on a sum over all funds and all months (i.e. these results are asset-weighted in the sense that they put more weight on larger funds). In Panel C and Panel D, we display similar results for deciles by book-to-market and past 12-months stock return. Deciles are formed each month.

Panel A: Holdings by size decile (1=small) – Fraction of shares outstanding

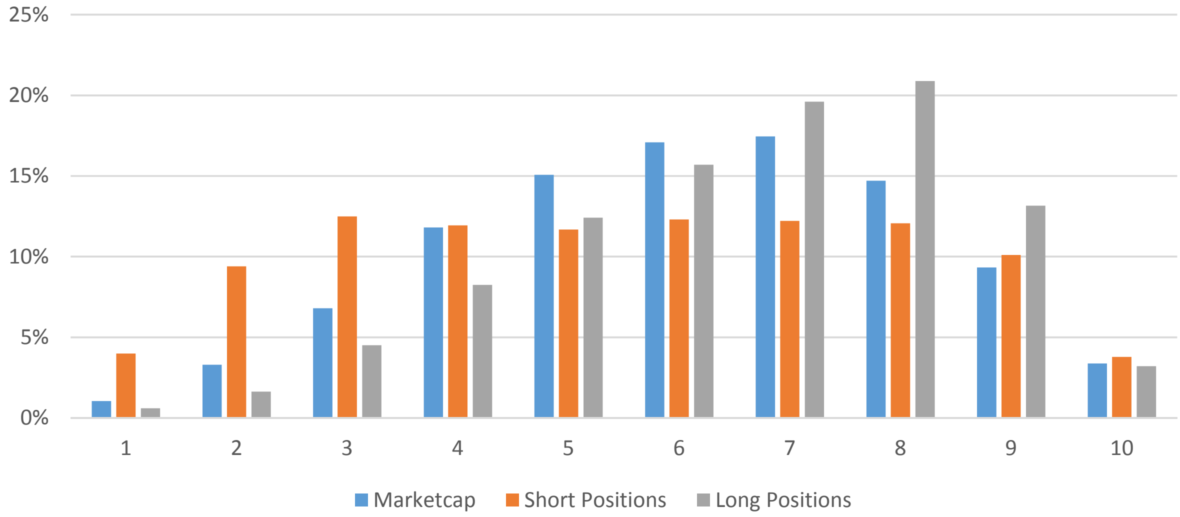


Panel B: Holdings by industry



Panel C: Holdings by book-to-market deciles (1=low book to market value)

Panel D: Holdings by past 12 month return (1=most negative return)



Section C: Robustness Checks

1) Profitability based on daily average holding-period returns

In Table 2 of the paper, we show that the opening of hedge fund positions predict returns—when cumulated over windows of 60 and 125 trading days (Panel A) or over the entire holding period (Panel B). In this robustness check, we examine (daily) *average* holding-period returns instead of cumulative holding period returns. We compute these returns from the last day of the opening order to the first day of the closing order (or the day the hedge fund leaves the sample). This is a conservative estimate because it excludes within-order profits, which on average are positive (unreported). As before, we regress these position-level returns on a dummy variable indicating whether it is a long position. We present the results in Table C.1 below. We find that long positions experience risk-adjusted returns that are on average 5bp higher per day (1.1% per year). This finding confirms that the hedge funds have investment skill in that their positions are profitable.

Table C.1: Robustness check for Table 2 (average holding-period returns)

This table examines whether daily average holding-period returns for long positions are more positive than average holding-period returns for short positions. The regression is run at the position level. The dependent variable is the average daily risk-adjusted return from the last day of the opening order to the first day of the closing order (or the day the hedge fund leaves the sample). We include fund fixed effects and month fixed effects (based on the month of the last day of the opening order). All standard errors are two-way clustered by stock and last date of the opening order. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Dependent Variable:	Daily Averages		
Sample	Benchmark-Adjusted Return	DGTW Return	4-Factor Alpha
	(1)	(2)	(3)
D(Long Position)	0.04***	0.05***	0.04***
	(3.87)	(4.49)	(3.80)
Observations	12452	9985	11231
Adjusted R^2	0.02	0.01	0.01
Fund Fixed Effects	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes

2) Different measures of market-wide financial constraints

In Table 7 of the paper, we show that hedge funds leave more money on the table when market-wide funding constraints tighten. We measure these constraints with the HKM intermediary risk factor and the TED spread. In Table C.2 below, we show that we obtain comparable results when we split the sample instead by *change in the VIX* or by *intermediary stock returns*. The VIX index is a measure of the implied volatility of S&P 500 index options, calculated and published by the Chicago Board Options Exchange (CBOE). Increases in the VIX are generally interpreted as reflecting an increase in risk aversion and tighter funding constraints. In Panel A and B, we show that the direction of a closing trade predicts future returns better after increases of the VIX over the past 5 (or 10) trading days. This finding suggests that funds exhibit more early position closures after increases in the VIX. Similarly, in Panel C and D, we show that there are more early position closures after negative intermediary stock returns. The intermediary stock returns, described in He, Manela and Kelly (2016), are value-weighted portfolio returns of all publicly-traded holding companies of primary dealer counterparties of the New York Federal Reserve. Negative returns signal that primary dealers have less capital and are more likely to tighten funding constraints for their client hedge funds. Both results suggest that hedge funds engage in more premature position closures when financial constraints tighten.

Table C.2: Robustness check for Table 7 (split by market-wide funding constraints)

This table examines whether returns following the closure of positions depend on changes in (market-wide) funding constraints. In Panels A and B, our proxy for funding constraints is the change in the VIX index over the prior 5 or 10 trading days. In Panels C and D, the proxy for funding constraints is the cumulative intermediary stock return, which is the value-weighted portfolio return of all publicly-traded holding companies of primary dealer counterparties of the New York Federal Reserve. These returns are available at http://apps.olin.wustl.edu/faculty/manela/data.html. The dependent variable is the cumulative return expressed in percent for 60 and 125 trading days following the last day of the order. Details on variable constructions can be found in Appendix A in the paper. We include fund fixed effects and month fixed effects (based on the month of the last day of the order). All standard errors are two-way clustered by stock and last date of order. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Panel A: Split by VIX change over prior 5 trading days

Dependent Variable:	Benchmark-adjusted Return t+1, t+125		DGTW Return t+1, t+125		4-Factor Alpha t+1, t+125	
Sample	Higher VIX	Lower VIX	Higher VIX	Lower VIX	Higher VIX	Lower VIX
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	1.73***	0.83	0.90	1.18**	0.93	0.83
	(2.65)	(1.39)	(1.51)	(2.05)	(1.49)	(1.49)
Observations	5959	5752	5345	5120	5905	5697
Adjusted R^2	0.10	0.11	0.07	0.05	0.05	0.05
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Split by VIX change over prior 10 trading days

Dependent Variable:	Benchmark-adjusted Return t+1, t+125		DGTW Return t+1, t+125		4-Factor Alpha t+1, t+125	
Sample	Higher VIX	Lower VIX	Higher VIX	Lower VIX	Higher VIX	Lower VIX
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	2.00***	0.48	1.24**	0.88	1.44**	0.29
	(3.11)	(0.79)	(2.08)	(1.48)	(2.39)	(0.49)
Observations	5956	5754	5320	5144	5915	5686
Adjusted R^2	0.11	0.09	0.07	0.06	0.05	0.04
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Panel C: Split by intermediary stock return over prior 5 trading days

Dependent Variable:	Benchmark-adjusted Return t+1, t+125		DGTW Return t+1, t+125		4-Factor Alpha t+1, t+125	
Sample	Negative Intermediary Return	Positive Intermediary Return	Negative Intermediary Return	Positive Intermediary Return	Negative Intermediary Return	Positive Intermediary Return
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	2.02***	0.71	1.07*	1.10*	1.48**	0.42
	(3.14)	(1.19)	(1.76)	(1.93)	(2.49)	(0.74)
Observations	5223	6488	4695	5770	5186	6416
Adjusted R^2	0.11	0.09	0.06	0.06	0.05	0.04
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Panel D: Split by intermediary stock return over prior 10 trading days

Dependent Variable:	Benchmark-adjusted Return t+1, t+125		DGTW Return t+1, t+125		4-Factor Alpha t+1, t+125	
Sample	Negative Intermediary Return	Positive Intermediary Return	Negative Intermediary Return	Positive Intermediary Return	Negative Intermediary Return	Positive Intermediary Return
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	1.83***	0.93	1.28**	0.99*	1.27**	0.66
	(2.85)	(1.56)	(2.05)	(1.78)	(2.12)	(1.15)
Observations	5151	6560	4605	5860	5112	6490
Adjusted R^2	0.10	0.10	0.05	0.07	0.05	0.05
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

3) Testing for backfill bias using DGTW returns and 4-factor alphas

In Table 11 in the paper, we show that funds do not exhibit statistically significantly different raw returns or benchmark-adjusted returns just after they enter or just before they leave the sample. In this subsection, we extend this test by studying DGTW returns and 4-factor alphas as dependent variables. We report the results in Table C.3 below. Generally, they confirm that returns are not significantly different at the beginning or the end of a fund being covered. The only exception is the positive and marginally significant coefficient for DGTW returns in the first 125 trading days. However, this result is likely random given that the comparable coefficients for alphas, benchmark-adjusted returns, and raw returns are far from being significant and sometimes even negative. In summary, these results do not indicate any evidence of backfill bias or sample selection.

Table C.3: Robustness check for Table 11 (testing for backfill bias and sample selection)

This table reports the same analysis as in Table 11 in the paper but using DGTW returns and 4-Factor Alphas as dependent variables. We examine whether hedge funds have different returns shortly after they enter or before they exit the sample. We run OLS regressions at the fund-date level. In Panel A, the dependent variable is the (position-weighted) average daily DGTW return of the funds' portfolio stocks. In Panel B, the dependent variable is the (position-weighted) average daily 4-factor alpha of the funds' portfolio stocks. The independent variables are dummy variables equal to one in the first (or last) 60 (or 120) days that the fund is reporting to Inalytics (and thus enters our sample). Details on variable constructions can be found in Appendix A. We include fund fixed effects and month fixed effects. All standard errors are clustered by date. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Panel A: DGTW Returns

Dependent Variable:	Daily Fund DGTW Return (in basis points)			
	(1)	(2)	(3)	(4)
D(First 60 days in sample)	0.48			
	(0.35)			
D(First 125 days in sample)		1.74^{*}		
		(1.66)		
D(Last 60 days in sample)			0.72	
			(0.42)	
D(Last 125 days in sample)				0.36
				(0.36)
Constant	1.04^{***}	0.85^{***}	1.03^{***}	1.02^{***}
	(3.88)	(3.00)	(3.83)	(3.62)
Observations	21257	21257	21257	21257
Adjusted R^2	0.00	0.00	0.00	0.00
Fund Fixed Effects	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes

Panel B: 4 Factor Alphas

Dependent Variable:	Daily Fund 4 Factor Alpha (in basis points)			
	(1)	(2)	(3)	(4)
D(First 60 days in sample)	-0.38			
	(-0.24)			
D(First 125 days in sample)		0.52		
		(0.43)		
D(Last 60 days in sample)			-0.61	
			(-0.30)	
D(Last 125 days in sample)				-0.77
				(-0.64)
Constant	1.33^{***}	1.24^{***}	1.34^{***}	1.40^{***}
	(3.12)	(2.84)	(3.15)	(3.26)
Observations	21260	21260	21260	21260
Adjusted R^2	0.00	0.00	0.00	0.00
Fund Fixed Effects	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes

4) Only using trades where holding and trade data agree

There are some inconsistencies between the holdings and trades data provided by Inalytics. Specifically, there are sometimes holding changes that are not accompanied by a matching trade. In the paper, we follow Inalytics' advice and assume that the holdings data are correct. That is, if there is a holding change but no recorded trade, we impute a trade that corresponds to the holding change. In Table C.4 below, we run a robustness check where we remove all trades inferred from the holdings data for which we do not have a trade that matches exactly in terms of stocks traded. In Panel A, we show robustness checks for the main specifications of Tables 2 and 3. In Panel B, we show robustness checks for the main specifications of Tables 4 to 6. Our results remain very similar, implying that they are not driven by errors in the Inalytics data.

Table C.4: Robustness check – Only trades where holding and trade data agree

This table shows a robustness check in which we remove all trades from our data for which the change in the holdings data does not exactly match the trade data. In Panel A, we show robustness checks for Tables 2 and 3. Regressions 1 and 2 are run on opening orders and provide robustness to Table 2 Panel A. Regressions 3 and 4 are run on closing orders and provide robustness to Table 2 Panel B. Regressions 5 and 6 are run on both closing and opening orders and provide robustness to Table 3 Panel C. In Panel B, we display robustness checks for the sample splits in Tables 4 to 6. Regressions 1 and 2 splits the sample based on the change in number of positions in the 5 days prior to the order. Regressions 3 and 4 split the sample based on whether the fund return in the 5 days prior to the order was positive. Regressions 5 and 6 split the sample by whether fund return volatility, measured as the sum of squared fund returns over the previous 20 trading days increased or decreased relative to the 20 trading days before that. Details on variable constructions can be found in Appendix A. We include fund fixed effects and month fixed effects (based on the month of the last day of the order), except for regressions 5 and 6 in Panel A which have fund×portfolio×month fixed effects instead. All standard errors are two-way clustered by stock and last date of order. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Panel A: Robustness for Tables 2-3

Sample:	Opening Orders		Closing Orders		Opening and Closing Orders	
Dependent Variable:	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Signed Benchmark-Adj. Ret. t+1, t+60	Signed Benchmark-Adj. Ret. t+1, t+125
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	1.86***	2.68***	0.73**	1.29**		
	(5.87)	(5.38)	(2.16)	(2.39)		
D(Position Opening)					0.55***	0.77***
					(2.93)	(2.72)
Observations	11836	11217	8813	8376	20649	19593
Adjusted R^2	0.06	0.09	0.07	0.09	0.13	0.15
Fund Fixed Effects	Yes	Yes	Yes	Yes	No	No
Month Fixed Effects	Yes	Yes	Yes	Yes	No	No
Fund×Portf.×Month F.E.	No	No	No	No	Yes	Yes

Panel B: Robustness for Tables 4-6 (sample splits)

Dependent Variable:	Benchmark-Adjusted Return t+1, t+125					
Sample	More Positions	Less Positions	Negative Fund Return	Positive Fund Return	Higher Volatility	Lower Volatility
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	1.42*	1.04	2.44***	0.58	1.59**	1.26*
	(1.81)	(1.44)	(2.95)	(0.88)	(2.15)	(1.69)
Observations	3947	4429	3753	4623	4050	4018
Adjusted R^2	0.09	0.10	0.13	0.07	0.10	0.10
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

5) Excluding trades with stock prices below $1

In our analysis, we remove any stocks below a stock price of USD 0.20. We choose this relatively low cut-off because international stocks often have low prices when converted to USD (without there being a rounding issue with the stock return in the local currency) and we want to exclude as few stocks as possible that are actually traded by our hedge funds. To show that our results are not driven by this low cut-off, we exclude all stocks with prices below USD 1 in the robustness check in Table C.5 below. Our results remain very similar.

Table C.5: Robustness check – Excluding trades with stock prices below $1

This table shows a robustness check in which we remove all trades of stocks with a price below USD 1. In Panel A, we show robustness checks for Tables 2 to 3. Regressions 1 and 2 are run on opening orders and provide robustness to Table 2 Panel A. Regressions 3 and 4 are run on closing orders and provide robustness to Table 2 Panel B. Regressions 5 and 6 are run on both closing and opening orders and provide robustness to Table 3 Panel C. In Panel B, we display robustness checks for the sample splits in Tables 4 to 6. Regressions 1 and 2 splits the sample based on the change in number of positions in the 5 days prior to the order. Regressions 3 and 4 split the sample based on whether the fund return in the 5 days prior to the order was positive. Regressions 5 and 6 split the sample by whether fund return volatility, measured as the sum of squared fund returns over the previous 20 trading days increased or decreased relative to the 20 trading days before that. Details on variable constructions can be found in Appendix A. We include fund fixed effects and month fixed effects (based on the month of the last day of the order), except for regressions 5 and 6 in Panel A which have fund×portfolio×month fixed effects instead. All standard errors are two-way clustered by stock and last date of order. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Panel A: Robustness for Tables 2-3

Sample:	Opening Orders		Closing Orders		Opening and Closing Orders	
Dependent Variable:	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Signed Benchmark-Adj. Ret. t+1, t+60	Signed Benchmark-Adj. Ret. t+1, t+125
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	1.89***	2.61***	0.70**	1.33***		
	(6.84)	(5.89)	(2.46)	(2.99)		
D(Position Opening)					0.56***	0.69***
					(3.84)	(3.04)
Observations	13483	12776	12143	11559	25626	24335
Adjusted R^2	0.06	0.08	0.07	0.09	0.13	0.15
Fund Fixed Effects	Yes	Yes	Yes	Yes	No	No
Month Fixed Effects	Yes	Yes	Yes	Yes	No	No
Fund×Portf.×Month F.E.	No	No	No	No	Yes	Yes

Panel B: Robustness for Tables 4-6 (sample splits)

Dependent Variable:	Benchmark-Adjusted Return t+1, t+125					
Sample	More Positions	Less Positions	Negative Fund Return	Positive Fund Return	Higher Volatility	Lower Volatility
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	2.10***	0.62	2.19***	0.71	1.86***	1.03*
	(3.25)	(1.07)	(3.22)	(1.33)	(3.02)	(1.69)
Observations	5153	6393	5095	6464	5597	5566
Adjusted R^2	0.10	0.09	0.11	0.08	0.10	0.10
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

6) Excluding trades of stocks outside stock universe

If funds trade stocks that are outside our stock universe used to compute DGTW returns (e.g. because they cannot be assigned to one of the regions or have no information on book value), we still include these trades in our sample (see Subsection A.3 of this internet appendix). For such trades, we only have access to benchmark-adjusted returns and alphas. In this robustness check, we limit our sample to stocks that are in our stock universe; that is, to stock trades for which we have benchmark-adjusted returns, alphas and DGTW returns. The results, reported in Table C.6 below, are very similar to those reported in the paper.

Table C.6: Robustness check – Excluding trades of stocks outside stock universe

This table shows a robustness check in which we remove all trades of stocks which are not within the stock universe (e.g., because region information or book value data is missing, etc.). In Panel A, we show robustness checks for Tables 2 to 3. Regressions 1 and 2 are run on opening orders and provide robustness to Table 2 Panel A. Regressions 3 and 4 are run on closing orders and provide robustness to Table 2 Panel B. Regressions 5 and 6 are run on both closing and opening orders and provide robustness to Table 3 Panel C. In Panel B, we display robustness checks for the sample splits in Tables 4 to 6. Regressions 1 and 2 splits the sample based on the change in number of positions in the 5 days prior to the order. Regressions 3 and 4 split the sample based on whether the fund return in the 5 days prior to the order was positive. Regressions 5 and 6 split the sample by whether fund return volatility, measured as the sum of squared fund returns over the previous 20 trading days increased or decreased relative to the 20 trading days before that. Details on variable constructions can be found in Appendix A. We include fund fixed effects and month fixed effects (based on the month of the last day of the order), except for regressions 5 and 6 in Panel A which have fund×portfolio×month fixed effects instead. All standard errors are two-way clustered by stock and last date of order. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Panel A: Robustness for Tables 2-3

Sample:	Opening Orders		Closing Orders		Opening and Closing Orders	
Dependent Variable:	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Signed Benchmark-Adj. Ret. t+1, t+60	Signed Benchmark-Adj. Ret. t+1, t+125
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	1.64***	2.17***	0.55*	1.28***		
	(5.63)	(4.67)	(1.84)	(2.75)		
D(Position Opening)					0.54***	0.47**
					(3.58)	(1.99)
Observations	11798	11331	11096	10573	22894	21904
Adjusted R^2	0.06	0.09	0.07	0.10	0.13	0.16
Fund Fixed Effects	Yes	Yes	Yes	Yes	No	No
Month Fixed Effects	Yes	Yes	Yes	Yes	No	No
Fund×Portf.×Month F.E.	No	No	No	No	Yes	Yes

Panel B: Robustness for Tables 4-6 (sample splits)

Dependent Variable:	Benchmark-Adjusted Return t+1, t+125					
Sample	More Positions	Less Positions	Negative Fund Return	Positive Fund Return	Higher Volatility	Lower Volatility
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	1.66**	0.87	2.15***	0.63	2.07***	0.83
	(2.50)	(1.42)	(3.02)	(1.13)	(3.23)	(1.32)
Observations	4743	5820	4665	5908	5100	5112
Adjusted R^2	0.10	0.10	0.12	0.08	0.10	0.11
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

7) Excluding return data from Analytics

For some trades, we cannot find matching return data in Datastream. In these cases, we use return data provided by Inalytics instead (see Subsection A.4 above). In Table C.7 below, we provide a robustness check using only return data from Datastream. The results remain very similar to those reported in the paper.

Table C.7: Robustness check – Exclude return data from Inalytics

This table shows a robustness check in which we remove all stock trades for which we do not have return data from Datastream. In Panel A, we show robustness checks for Tables 2 to 3. Regressions 1 and 2 are run on opening orders and provide robustness to Table 2 Panel A. Regressions 3 and 4 are run on closing orders and provide robustness to Table 2 Panel B. Regressions 5 and 6 are run on both closing and opening orders and provide robustness to Table 3 Panel C. In Panel B, we display robustness checks for the sample splits in Tables 4 to 6. Regressions 1 and 2 splits the sample based on the change in number of positions in the 5 days prior to the order. Regressions 3 and 4 split the sample based on whether the fund return in the 5 days prior to the order was positive. Regressions 5 and 6 split the sample by whether fund return volatility, measured as the sum of squared fund returns over the previous 20 trading days increased or decreased relative to the 20 trading days before that. Details on variable constructions can be found in Appendix A. We include fund fixed effects and month fixed effects (based on the month of the last day of the order), except for regressions 5 and 6 in Panel A which have fund×portfolio×month fixed effects instead. All standard errors are two-way clustered by stock and last date of order. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Panel A: Robustness for Tables 2-3

Sample:	Opening Orders		Closing Orders		Opening and Closing Orders	
Dependent Variable:	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Signed Benchmark-Adj. Ret. t+1, t+60	Signed Benchmark-Adj. Ret. t+1, t+125
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	1.76***	2.41***	0.68**	1.46***		
	(6.13)	(5.35)	(2.33)	(3.19)		
D(Position Opening)					0.53***	0.54**
					(3.56)	(2.38)
Observations	12587	12093	11759	11192	24346	23285
Adjusted R^2	0.06	0.09	0.07	0.10	0.14	0.16
Fund Fixed Effects	Yes	Yes	Yes	Yes	No	No
Month Fixed Effects	Yes	Yes	Yes	Yes	No	No
Fund×Portf.×Month F.E.	No	No	No	No	Yes	Yes

Panel B: Robustness for Tables 4-6 (sample splits)

Dependent Variable:	Benchmark-Adjusted Return t+1, t+125					
Sample	More Positions	Less Positions	Negative Fund Return	Positive Fund Return	Higher Volatility	Lower Volatility
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	2.17***	0.82	2.41***	0.72	2.12***	1.08*
	(3.27)	(1.40)	(3.45)	(1.33)	(3.35)	(1.75)
Observations	5022	6157	4919	6273	5393	5415
Adjusted R^2	0.11	0.10	0.12	0.08	0.10	0.11
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

8) Excluding trades around mergers events

A common hedge fund strategy is to engage in merger arbitrage; i.e., purchasing the target and short selling the acquirer of an announced stock merger, thereby betting on its completion. We show in Subsection C.3 below that our funds almost never engage in merger arbitrage. Nonetheless, one may wonder whether our results can be confounded by merger events. In this subsection, we therefore provide a robustness check in which we exclude the days around a merger event. Specifically, we exclude from our sample all stock-days for both the acquirer and the target starting from a week before the announcement of a merger until one week after the merger is either completed or withdrawn. Table C.8 below shows that our results remain very similar after excluding these observations.

Table C.8: Robustness check – Exclude days around mergers

This table shows a robustness check in which we remove all stock-days for both the target and the acquirer stock starting from 7 days before the announcement of the merger to 7 days after the merger is either completed or withdrawn. If we do not have information on merger completion, we assume that the merger completes 30 days after its announcement. In Panel A, we show robustness checks for Tables 2 to 3. Regressions 1 and 2 are run on opening orders and provide robustness to Table 2 Panel A. Regressions 3 and 4 are run on closing orders and provide robustness to Table 2 Panel B. Regressions 5 and 6 are run on both closing and opening orders and provide robustness to Table 3 Panel C. In Panel B, we display robustness checks for the sample splits in Tables 4 to 6. Regressions 1 and 2 splits the sample based on the change in number of positions in the 5 days prior to the order. Regressions 3 and 4 split the sample based on whether the fund return in the 5 days prior to the order was positive. Regressions 5 and 6 split the sample by whether fund return volatility, measured as the sum of squared fund returns over the previous 20 trading days increased or decreased relative to the 20 trading days before that. Details on variable constructions can be found in Appendix A. We include fund fixed effects and month fixed effects (based on the month of the last day of the order), except for regressions 5 and 6 in Panel A which have fund×portfolio×month fixed effects instead. All standard errors are two-way clustered by stock and last date of order. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Panel A: Robustness for Tables 2-3

Sample:	Opening Orders		Closing Orders		Opening and Closing Orders	
Dependent Variable:	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Signed Benchmark-Adj. Ret. t+1, t+60	Signed Benchmark-Adj. Ret. t+1, t+125
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	1.87***	2.51***	0.69**	1.30***		
	(6.60)	(5.47)	(2.41)	(2.94)		
D(Position Opening)					0.55***	0.67***
					(3.59)	(2.91)
Observations	13222	12554	11906	11379	25128	23933
Adjusted R^2	0.06	0.09	0.07	0.10	0.13	0.16
Fund Fixed Effects	Yes	Yes	Yes	Yes	No	No
Month Fixed Effects	Yes	Yes	Yes	Yes	No	No
Fund×Portf.×Month F.E.	No	No	No	No	Yes	Yes

Panel B: Robustness for Tables 4-6 (sample splits)

Dependent Variable:	Benchmark-Adjusted Return t+1, t+125					
Sample	More Positions	Less Positions	Negative Fund Return	Positive Fund Return	Higher Volatility	Lower Volatility
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	2.04***	0.62	1.97***	0.79	1.79***	1.01
	(3.14)	(1.08)	(2.91)	(1.47)	(2.97)	(1.64)
Observations	5066	6301	4999	6380	5494	5496
Adjusted R^2	0.10	0.10	0.11	0.09	0.10	0.11
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

9) Excluding stocks in EME region or without region assignment

The EME region includes countries such as Thailand and Venezuela, which are fairly different. In addition, there are observations for which we do not have access to regional factors and therefore compute 4-factor alphas using global factors. One concern may be that using a broad region such as EME or no regional factors at all, may lead to insufficient risk-adjustment of returns. In this subsection, we therefore provide a robustness check in which we exclude all stocks in the EME region and all stocks without region assignment. As show in Table C.9 below, the results remain similar, suggesting that they are not driven by insufficient risk-adjustment of returns.

Table C.9: Robustness check – Exclude stocks in EME region or without region assignment

This table shows a robustness check in which we remove all stocks in the EME region or that we cannot assign to a region. In Panel A, we show robustness checks for Tables 2 to 3. Regressions 1 and 2 are run on opening orders and provide robustness to Table 2 Panel A. Regressions 3 and 4 are run on closing orders and provide robustness to Table 2 Panel B. Regressions 5 and 6 are run on both closing and opening orders and provide robustness to Table 3 Panel C. In Panel B, we display robustness checks for the sample splits in Tables 4 to 6. Regressions 1 and 2 splits the sample based on the change in number of positions in the 5 days prior to the order. Regressions 3 and 4 split the sample based on whether the fund return in the 5 days prior to the order was positive. Regressions 5 and 6 split the sample by whether fund return volatility, measured as the sum of squared fund returns over the previous 20 trading days increased or decreased relative to the 20 trading days before that. Details on variable constructions can be found in Appendix A. We include fund fixed effects and month fixed effects (based on the month of the last day of the order), except for regressions 5 and 6 in Panel A which have fund×portfolio×month fixed effects instead. All standard errors are two-way clustered by stock and last date of order. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Panel A: Robustness for Tables 2-3

Sample:	Opening Orders		Closing Orders		Opening and Closing Orders	
Dependent Variable:	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Benchmark-Adj. Ret. t+1, t+60	Benchmark-Adj. Ret. t+1, t+125	Signed Benchmark-Adj. Ret. t+1, t+60	Signed Benchmark-Adj. Ret. t+1, t+125
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	1.87^{***}	2.51^{***}	0.69^{**}	1.30^{***}		
	(6.60)	(5.47)	(2.41)	(2.94)		
D(Position Opening)					0.55^{***}	0.67^{***}
					(3.59)	(2.91)
Observations	13222	12554	11906	11379	25128	23933
Adjusted R^2	0.06	0.09	0.07	0.10	0.13	0.16
Fund Fixed Effects	Yes	Yes	Yes	Yes	No	No
Month Fixed Effects	Yes	Yes	Yes	Yes	No	No
Fund×Portf.×Month F.E.	No	No	No	No	Yes	Yes

Panel B: Robustness for Tables 4-6 (sample splits)

Dependent Variable:	Benchmark-Adjusted Return t+1, t+125					
Sample	More Positions	Less Positions	Negative Fund Return	Positive Fund Return	Higher Volatility	Lower Volatility
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	2.04^{***}	0.62	1.97^{***}	0.79	1.79^{***}	1.01
	(3.14)	(1.08)	(2.91)	(1.47)	(2.97)	(1.64)
Observations	5066	6301	4999	6380	5494	5496
Adjusted R^2	0.10	0.10	0.11	0.09	0.10	0.11
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

10) Including within-order returns

In our main analysis, we look at returns measured from the closing price of the last day of the order, so as to not confound post-trade returns with any profits or losses earned during the trading day. In Table C.10 below, we show that our results are unaffected when we instead include within-order returns. The following example illustrates how we include within-order returns: A fund enters a new long position by buying 100 stocks at \$55 on day 1 and 100 stocks at \$60 on day 2. In this robustness check, we then compute returns relative to the average purchase price, in this case \$57.5 (instead of using the closing price on day 2, as was done in our main analysis).

Table C.10: Robustness check – Include within-order returns

This table shows a robustness check in which we measure returns starting from the average transaction price (instead of from the closing price on the last day of the order), thereby including within-order returns. In Panel A, we show robustness checks for Tables 2 to 3. Regressions 1 and 2 are run on opening orders and provide robustness to Table 2 Panel A. Regressions 3 and 4 are run on closing orders and provide robustness to Table 2 Panel B. Regressions 5 and 6 are run on both closing and opening orders and provide robustness to Table 3 Panel C. In Panel B, we display robustness checks for the sample splits in Tables 4 to 6. Regressions 1 and 2 splits the sample based on the change in number of positions in the 5 days prior to the order. Regressions 3 and 4 split the sample based on whether the fund return in the 5 days prior to the order was positive. Regressions 5 and 6 split the sample by whether fund return volatility, measured as the sum of squared fund returns over the previous 20 trading days increased or decreased relative to the 20 trading days before that. Details on variable constructions can be found in Appendix A. We include fund fixed effects and month fixed effects (based on the month of the last day of the order), except for regressions 5 and 6 in Panel A which have fund×portfolio×month fixed effects instead. All standard errors are two-way clustered by stock and last date of order. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Panel A: Robustness for Tables 2-3

Sample:	Opening Orders		Closing Orders		Opening and Closing Orders	
Dependent Variable:	Benchmark-Adj. Ret. t_{trade}, t+60	Benchmark-Adj. Ret. t_{trade}, t+125	Benchmark-Adj. Ret. t_{trade}, t+60	Benchmark-Adj. Ret. t_{trade}, t+125	Signed Benchmark-Adj. Ret. t_{trade}, t+60	Signed Benchmark-Adj. Ret. t_{trade}, t+125
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	2.17***	2.74***	0.70**	1.33***		
	(7.61)	(6.06)	(2.48)	(2.96)		
D(Position Opening)					0.74***	0.82***
					(4.90)	(3.53)
Observations	13758	13046	12432	11839	26190	24885
Adjusted R^2	0.06	0.09	0.07	0.10	0.13	0.16
Fund Fixed Effects	Yes	Yes	Yes	Yes	No	No
Month Fixed Effects	Yes	Yes	Yes	Yes	No	No
Fund×Portf.×Month F.E.	No	No	No	No	Yes	Yes

Panel B: Robustness for Tables 4-6 (sample splits)

Dependent Variable:	Benchmark-Adjusted Return t_{trade}, t+125					
Sample	More Positions	Less Positions	Negative Fund Return	Positive Fund Return	Higher Volatility	Lower Volatility
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	2.04***	0.66	2.30***	0.58	1.84***	0.99
	(3.12)	(1.13)	(3.35)	(1.09)	(2.94)	(1.63)
Observations	5279	6547	5213	6626	5718	5721
Adjusted R^2	0.10	0.10	0.11	0.09	0.10	0.11
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Section D: Additional Results

1) Do funds hedge their industry exposure?

In this subsection, we study whether our funds engage in industry hedging; i.e., whether they attempt to limit their industry exposure by going short and long different stocks of the same industry. To this end, we group stock positions for every fund and every day by 2-digit SIC codes. For each newly opened stock position, we create the dummy variable *D(Short exposure by 2-digit SIC code)* that equals one if the other holdings of the fund in the same industry on the previous day are more short than long (in terms of dollar value). We then regress *D(Long Position)* on *D(Long exposure by 2-digit SIC code).* We report the results in regression 1 of Table C.2 below. If hedge funds engaged in industry hedging, we would expect significant negative coefficients. Instead, the coefficient is positive but very close to zero and statistically insignificant. Regression 2 documents similar results for the 49 Fama-French industry classification, even though now the coefficient is slightly negative. These findings suggest that hedge funds do not hedge their industry exposure. As shown in regression 3, there seems to be no hedging by DGTW portfolios (quintiles based on size, market-to-book and past returns) either. Finally, we also do not find any significant results when interacting these industry and DGTW groups with the investment region in regressions 4-6. Taken together, these findings suggest that stock positions are not meant to hedge away existing risk exposures to specific industries or stock characteristic.

Table D.1: Industry hedging

This table examines whether hedge funds concentrate their short and long positions in certain industries and DGTW portfolios. The sample includes only opening orders. The dependent variable is *D(Long Position)*, which is a dummy variable equal to one if the newly opened position is a long position and zero if it is a short position. The explanatory variable is *D(Long exposure by group)*, which is a dummy variable equal to one if the *existing* holdings of the fund in this group are more long than short (by dollar value) on the day before the first day of the order. We use 6 groups: 2-digit SIC industries, 49 Fama-French industries, DGTW portfolios (quintiles by size, market-to-book and past returns) and each of these groups interacted with the investment region. Details on variable constructions can be found in Appendix A of the paper. We include fund fixed effects, month fixed effects and group fixed effects based on the specific group. All standard errors are two-way clustered by group and date. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Dependent Variable:	D(Long Position)					
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long exposure by 2-digit SIC code)	0.01					
	(0.18)					
D(Long exposure by 49 FF industry)		-0.03				
		(-0.29)				
D(Long exposure by DGTW portfolio)			0.02			
			(0.48)			
D(Long exposure by 2-digit SIC code * region)				-0.02		
				(-0.29)		
D(Long exposure by 49 FF industry * region)					-0.06	
					(-0.67)	
D(Long exposure by DGTW portfolio * region)						-0.01
						(-0.29)
Observations	11479	11981	9030	9288	9635	8173
Adjusted R^2	0.08	0.07	0.06	0.09	0.07	0.07
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

2) Pairs trading and merger arbitrage

In this subsection, we investigate whether our sample hedge funds systematically engage in merger arbitrage or pairs trading. Both of these popular arbitrage strategies involve pairs of long and short trades; and thus such stock trades could hardly be considered as independent. Merger arbitrageurs usually purchase the target and short sell the acquirer, thereby betting on completion of the merger. This leads us to examine whether our hedge funds tend to hold positions of opposite directions in the acquirer and target between the announcement and the completion of an acquisition. For this purpose, we collect from SDC Platinum all acquisitions where both the acquirer and the target are publicly listed. We obtain a total of 17,593 merger events. In only 96 of those, one of our sample hedge funds holds a stake in both the acquirer and the target and in 60% of these cases, the positions have the same direction (i.e., they are not long-short). Restricting the analysis to stock positions that are established in the two weeks after a merger announcement, we find that there is only 1 merger event in which a sample fund opened both a long position in the target and a short position in the acquirer.

Turning to pairs trading, we note that this strategy consists of finding two highly correlated stocks and then going long (short) the relatively under- (over-)valued stock of the pair. As such, pairs trading should show up in the data as the opening of both a long and a short position in a pair of highly correlated stocks. We use this insight to test for the extent of pairs trading in our trading data. Specifically, we consider all position pairs that each fund opens on the same day and compute the correlation of the pairs' stocks over the past 250 trading days. We then check whether pairs with a high return correlation are indeed more likely to be opened as a long-short trade (that is, whether hedge funds go long in one stock of the pair and short in the other). This test takes the form of regressing *D(Long-Short)*, a dummy variable equal to one for long-short trades, on the correlation coefficient.

The results are reported in Table D.2 below. We find a significantly negative coefficient, showing that funds are less likely to open long-short positions if the two stocks of the pair are more strongly correlated. We find similar results when we replace the correlation coefficient by dummy variables that flag correlations above 50% or 75%; that is, for stock pairs that are particularly well-suited for a pairs trading strategy. Taken together, these results suggest that the funds in our sample do not engage in merger arbitrage or pairs trading.

Table D.2: Pairs trading

This table examines whether hedge funds are more likely to take positions with opposite direction in highly correlated stocks. The sample consists of all stock pairs in which a hedge fund initiates a position on the same day. The dependent variable is *D(Long-Short)*, which is a dummy variable equal to one if the opened positions go in opposite directions, i.e. one is a long and the other a short position. The explanatory variables are *Correlation Coefficient*, which is the correlation coefficient between the two stocks of the pair over the previous 250 trading days, as well as *D(Correlation Coefficent > 0.50)* and *D(Correlation Coefficient > 0.75)*, which are dummy variables equal to one if the correlation coefficient is above 50% or 75%, respectively. We include fund fixed effects and month fixed effects. We cluster standard errors by date. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Dependent Variable:	D(Long-Short)		
	(1)	(2)	(3)
Correlation Coefficient	-0.17***		
	(-4.00)		
D(Correlation Coefficient > 0.50)		-0.05***	
		(-4.52)	
D(Correlation Coefficient > 0.75)			-0.08**
			(-2.56)
Observations	31509	31509	31509
Adjusted R^2	0.05	0.05	0.05
Fund Fixed Effects	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes

3) Can rebalancing explain our results?

As we argue in the paper, rebalancing should not affect closing trades because by definition a rebalancing trade never closes a position completely.

In this subsection, we present an additional test to confirm this argument. Specifically, we examine a sample split for post-closure returns by the underlying stock's return over the prior 10 trading days. If one were to believe that rebalancing trades can close a position, one would expect such rebalancing trades to be more likely to occur after a positive return because positive returns increase the size of a position. Thus, if the alpha following closing orders was explained by rebalancing, we would expect a larger alpha if the closing happens after a positive stock return. Our regression setup is identical to the sample splits provided in the paper (Tables 4-7), except that we now split by the closing order's prior stock return. The results are presented below in Table D.3. We find very similar returns following closing orders after positive and negative stock returns. If anything, hedge funds seem to leave slightly more money on the table when they close positions after negative stock returns, which is the exact opposite of what we one would expect if closing orders were due to rebalancing. This finding confirms our argument that portfolio rebalancing cannot explain early position closures.

Table D.3: Returns after the closure of positions – Split by past stock return

This table examines whether returns following the closure of positions depend on the stocks prior return. We run the same regression as in Table 3 but split the sample by whether the stock had a positive or negative return in the 10 trading days prior to the first day of the order. We regress average returns following the order on a dummy variable whether the order is related to a short position (and is thus a short buy). The dependent variable is the cumulative return expressed in percent for 60 and 125 trading days following the last day of the order. Details on variable constructions can be found in Appendix A.1. We include fund fixed effects and month fixed effects (based on the month of the last day of the order). All standard errors are two-way clustered by stock and last date of order. We report t-statistics below the coefficients in parenthesis. ***, **, * indicate significance at the 1%, 5% and 10% level.

Dependent Variable:	Benchmark-Adjusted Return t+1, t+125		DGTW Return t+1, t+125		4-Factor Alpha t+1, t+125	
Sample	Positive Stock Return	Negative Stock Return	Positive Stock Return	Negative Stock Return	Positive Stock Return	Negative Stock Return
	(1)	(2)	(3)	(4)	(5)	(6)
D(Long Position)	-1.09*	-1.59**	-1.03*	-1.00*	-0.97*	-1.01*
	(-1.87)	(-2.56)	(-1.87)	(-1.76)	(-1.74)	(-1.66)
Observations	5671	5401	5348	5119	5675	5395
Adjusted R^2	0.09	0.12	0.06	0.06	0.05	0.06
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

4) Are our hedge funds subject to the disposition effect?

As we argue in the paper, the disposition effect leads traders to close their winning positions too early and may thus be an explanation for early position closures. We therefore examine in this subsection whether our funds are subject to the disposition effect. We follow Odean (1998) and compute the proportion of gains (losses) realized as:

$$Proportion\ of\ Gains\ Realized\ (PGR) = \frac{Realized\ Gains}{Realized\ Gains + Paper\ Gains}$$

$$Proportion\ of\ Losses\ Realized\ (PLR) = \frac{Realized\ Losses}{Realized\ Losses + Paper\ Losses}$$

We define gains (losses) as stocks that have experienced a positive (negative) signed stock return since the last day of the opening order. We define a position as realized (i.e. closed) on the last day of the closing order. Following Odean (1998), we include only fund-days on which the fund closes at least one position. We compute averages of PGR and PLR on these days and test whether they are significantly different. The results are reported below in Table D.4. In Column 1, we include all months of the sample. The proportion of losses realized is significantly larger than the proportion of gains realized. This result is the opposite of Odean (1998) and shows that our funds are not subject to the disposition effect but rather are more likely to close losing positions, i.e. they engage in stop-loss trading. In Columns 2 to 4, we show January and December separately because they may be affected by tax considerations. In all specifications, PLR is larger than PGR (although insignificantly so in December). These findings suggest that the hedge funds are not subject to the disposition effect.

Table D.4: Proportion of Gains and Losses Realized

This table examines whether hedge funds are more likely to close winning or losing positions. For each day in which a fund closes at least one position, we compute the proportion of gains and losses realized. We define gains (losses) as stocks that have experienced a positive (negative) signed stock return since the last day of the opening order. We define a position as realized (i.e. closed) on the last day of the closing order. In Column 1, we include days from all months. In Columns 2-4, we include only days from certain months. We report the difference between the proportions at the bottom of the table and compute a t-statistic for the difference. ***, **, * indicate significance at the 1%, 5% and 10% level.

Sample	All month	Only January	Only December	February - November
	(1)	(2)	(3)	(4)
Proportion of Losses Realized (PLR)	4.61	4.34	5.21	4.59
Proportion of Gains Realized (PGR)	4.06	3.77	5.10	4.01
Difference in Proportions (PLR – PGR)	0.55***	0.57	0.11	0.58***
(t-statistic)	(3.80)	(1.35)	(0.16)	(3.75)

References

Busse, Jeffrey A., Amit Goyal, and Sunil Wahal, Investing in a Global World, *Review of Finance* 18, 561-590.

Carhart, Mark M., 1997, On Persistence in Mutual Fund Performance, *Journal of Finance* 52, 57-82.

Daniel, Kent, Mark Grinblatt, Sheridan Titman and Russ Wermers, 1997, Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, *Journal of Finance 52*, 1035-1058.

Fama, Eugene F. and Kenneth R. French, 2012, Size, Value, and Momentum in International Stock Returns, *Journal of Financial Economics* 103, 457-472.

Gompers, Paul A., and Andrew Metrick, 2001, Leverage, Institutional Investors and Equity Prices, *Quarterly Journal of Economics* 106, 229-259.

Karolyi, Andrew, and Ying Wu, 2014, Size, Value, and Momentum in International Stock Returns: A New Partial-Segmentation Approach, *Working Paper*.

Ince, Ozgur S., and R. B. Porter, 2004, Individual Equity Return Data from Thomson Datastream: Handle with Care! *SSRN eLibrary*.

Odean, Terrance, 1998, Are Investors Reluctant to Realize Their Losses? *The Journal of Finance* 53, 1775-1798.