A Appendices

A.1 Data construction: additional details

The primary source of our analysis is Inalytics' holdings data and changes in holdings. We apply two primary filters to select the set of portfolios to include in our analysis. First, we drop portfolios for which daily trading data are unavailable or appear to be incomplete.[57] Second, we exclude funds that do not have a sufficient fraction (at least 80 percent) of portfolio holdings which could be reliably matched with CRSP or Datastream. In demonstrating the robustness of our results, we perform the analyses using data for larger stocks (using market cap information from Worldscope) and those with higher trading volume. These markets arguably have better price discovery and higher match rates with CRSP/Datastream.

After we clean the holdings data, we convert all the prices into USD using exchange rates from Datastream. To ensure accuracy in exchange rates, we compare the exchange rate in Datastream with two other sources of exchange rates from Compustat and Inalytics. In the event of a discrepancy, we pick the two out of three that are the same, and this procedure takes care of discrepancy in all cases. We then augment the holdings data by merging in external prices series and returns from CRSP (US stocks), Datastream (International stocks) and Inalytics' provided price series in this order. The external price series allow us to compute the market value of each holding precisely. There are instances where the market value of a stock (likely due to a measurement error in price/quantity) seems implausibly high, so we employ an iterative weight cleaning algorithm to eliminate these positions from the analysis. We provide additional details about these steps below.

We begin by outlining the key steps of our data cleaning procedure:

1. **Cleaning identifiers**: Inalytics has four main types of identifiers for stocks: SEDOL, ISIN CUSIP, and LOCAL. For the first three types of identifiers, they are distinguishable by the number of digits. SEDOL has 6-7 digits, CUSIP has 8-9 digits, and ISIN has 12 digits. In a few instances one type of identifier is mislabeled by the clients, so we correct them according to the number of digits.

[57] Trades are sometimes imputed at month-end because Inalytics receives portfolio snapshots in adjacent months which do not fully match the aggregated the trade data. This requires a reconciliation process. Because of this, we exclude funds that have a large fraction of trades occurring at the end of each month.

2. **Merging in liquidated stocks with holdings data**: There are instances when a fund completely closes a position, so a stock disappears from the holdings data. Since our main trade measure is computed from the change in stock's holding, a position-closing trade will not be observed in the holdings. To do so, we first measure the minimum and maximum dates of each fund. Then, we locate the instances where the stocks disappears from the portfolio between the start and end dates. We then append those stocks back to the holdings data.

3. **Dropping portfolios without daily trades**: Some of the portfolios in the dataset do not receive daily time-stamped trade data. In these cases, only monthly holdings are reported and trades are imputed at the end of the month. To filter out these portfolios, we count the fraction of trades after the 27th of the month for each fund. If a fraction of trades after the 27th for a fund is over 50% or missing (in case of no trades observed), we drop the portfolio from the analysis sample. In addition, Inalytics independently provided a list of these portfolios from their internal records, essentially all of which were filtered out by this criterion. We also remove these manually flagged portfolios.

Next, we discuss some potential issues related to measurement errors in the price data. Where available, we use external price series from CRSP and Datastream, and we supplement the remainder with data provided by Inalytics. Inalytics relies on multiple data vendors, including MSCI and Thompson Reuters, as well as clients themselves, for price series in the holdings data. In some cases, reported prices are overstated, which would lead us to incorrectly characterize portfolio weights and potentially introduce measurement errors in various counterfactual return calculations. We rely on our external price series as the primary measure for a price when computing returns and portfolio weights throughout the analysis, only relying on Inalytics where we otherwise would lack a security price. Our performance measures only include stocks which have a valid price series for the full holding period.

When we compute cumulative returns for purposes of evaluating trading performance, we winsorize extremely small and large return realizations, some of which may be due to measurement errors in the price data. To mitigate the effect of the extreme returns when computing the average returns, we winsorize returns in the holdings dataset across all measures (raw and factor neutral) before forming portfolios. In our baseline results that we present here, we employ two winsorizing thresholds. First, we winsorize the cumulative return measures on each date across all positions at 0.1% on either tail. As an additional precaution, we winsorize large positive returns in the whole sample at the 99.99% threshold on the right tail of the distribution for raw returns and 0.01% on either tail for factor neutral returns. The rationale is that factor neutral returns can have also have extreme negative returns after adjusting for risk, so it is necessary to winsorize on both tails for risk-adjusted returns measures. We have also considered larger thresholds for winsorizing such as 0.3%, 0.5%, and 1% and obtain similar results.

In a handful of cases (e.g., because a stock split has led to an incorrectly high market value), the market value of a single position appears to be extremely large relative to the rest of the portfolio, which is indicative of a likely measurement issue. In order to flag situations when one errant price could cause our estimates of portfolio weights to be substantially biased, we employ an iterative procedure to drop potentially problematic positions. In essence, we look for situations where the entire portfolio is concentrated in a single, extremely large position. For these purposes, we compute the market value of a position as the minimum of raw Inalytics price and raw external prices times the quantity of stocks, and then compute the position-level weight by dividing through by the dollar value of all positions. With these weights in hand, the procedure proceeds as follows. First, we compute the first three largest weights at a portfolio-date level. We then compute two measures 1) the difference between weights of the largest and second largest-held stocks and 2) the difference between weights of the second and third largest-held stocks. If the first difference minus the second difference is over 15%, the largest weight is over 10% and the second difference is less than 5%, we flag the stock with the largest weight, and exclude it from the analysis and the weight calculation. We then recompute stocks' weights after the largest-held stocks are dropped and repeat the procedure to flag other stocks with unusually high weight in a portfolio. We repeat this algorithm until there is no stock with an unusually large weight in portfolios. This iterative weight-dropping algorithm finishes in 5 runs. There are 57,982 stock-date observations to be excluded from weight calculation. 84.3M observations (94.12%) in the holdings data have no weight errors. The first run of this algorithm cleans up weights for 4.1M observations (4.62%) in the holdings data. After five runs of this algorithm, whereby we exclude five stocks at most, 99.86% of holding observations have no weight problems. There are two portfolios for which this procedure still indicates the presence of a handful of extremely concentrated positions but these two portfolios only made up 39,128 out of 89M observations.

We use the Reuters Worldscope database of equity characteristics for information other than daily prices and volume. When we need to perform filters using market cap, compute characteristics quintiles for DGTW-style adjustments, or perform any other analysis based on the fundamentals (e.g., book-to-market ratio) of the securities, we use this merged sample, which matches 80% of our full sample. When we perform for liquidity adjustments, we merge in daily volume from Datastream, which has data from global exchanges converted to USD. Datastream volume covers 83% of our daily holdings data, and we use this subsample when constructing liquidity-based subsamples or dropping securities below a certain Amihud measure threshold.

The Amihud measure is calculated on a monthly basis using daily dollar volume (in USD) from Datastream and returns from our sample of security returns (described earlier in the appendix). We use the prior month's worth of daily data to construct the illiquidity measure, which is the average of the absolute value of daily returns over that period divided by the dollar volume. Because we are interested in keeping the most liquid stocks in the counterfactual, we use the set of securities for which this measure is available in our matched sample when calculating a 5% breakpoint.

In Table 2, where we analyze whether our buying and selling skill measures are associated with the excess return of the fund as a whole, we compute the Buy - Hold and Hold - Sell portfolios as we do in the rest of the paper, and subsequently take averages of those measures over all dates. Hold - Benchmark is the value of the hold portfolio return (over some future window) less the benchmark return over the equivalent window. The buy, sell, and hold portfolios are adjusted for factor exposures to value, momentum, size, and market. Fund returns are similarly calculated on a rolling basis and then averaged over the full period. Fund returns are not adjusted for factor exposures, and neither are benchmarks. Benchmarks are self-designated by PMs/clients when initiating the Inalytics monitoring service. Inalytics provides daily benchmark-adjusted return series, which we compound over the relevant period for purposes of the regression.

We construct a measure of net flows in and out of a given strategy at various horizons. The flow is estimated as the remainder of total assets under management not explained by the prior returns of the fund. In other words, if we had 100 dollars in a fund, we observe a period return of 20%, and the subsequent period had 140 dollars in the fund, we consider the $20 remainder as the net flow of additional AUM injected over the period. Note that changes in net cash positions will look like flows according to our measure, since we do not observe cash positions in the holdings data.

A.2 Additional descriptive statistics

In order to better understand the sample, we collect information on the portfolios by hand. Because we only observe names of the portfolios as they are entered into Inalytics' system (not, for example, the specific traceable vehicles that are readily merged with external data), we are only able to identify a subset of manager and portfolio names. Often these portfolios are associated with specific separately managed accounts which are constructed with custom characteristics for clients, and thus may not have a related public strategy. Portfolios can have names which lack obvious identifying characteristics, making them impossible to match to a specific firm or strategy.

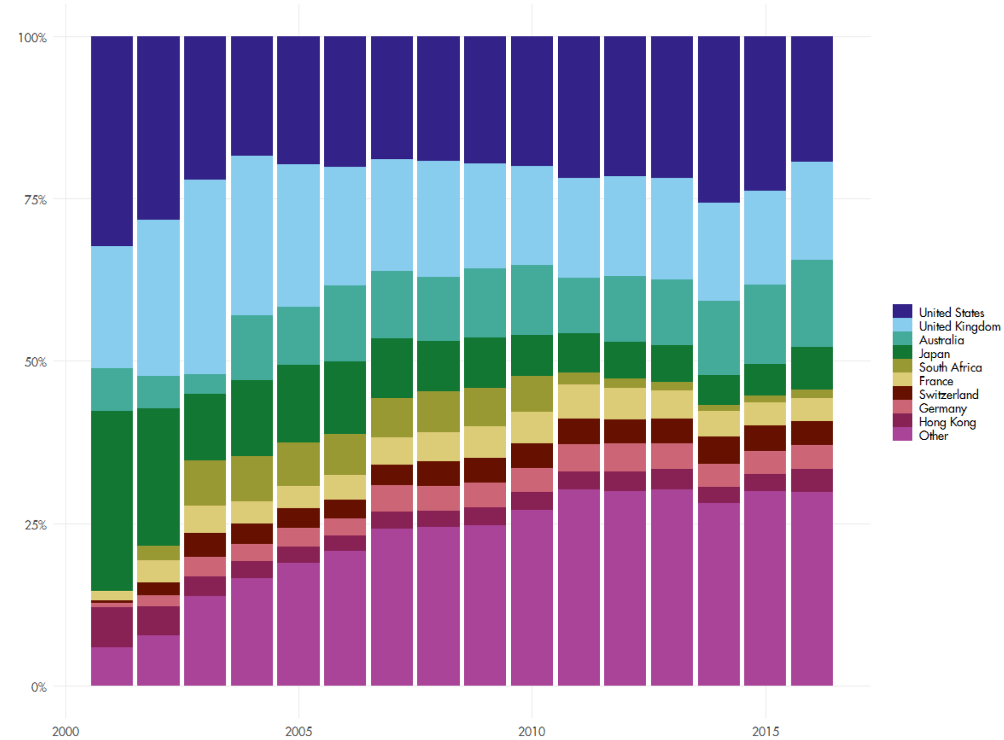
Where possible (70% of the sample), we identify the management company, and hand-classify them as Fundamental, Quantitative, Other, or some combination based on the way that they describe themselves on their website. We find only 22 managers in the sample are purely quantitative, with the remainder being primarily fundamental (sometimes with quantitative filters performed before an analyst reviews a universe). The vast majority describe their investment process as being built around custom research from internal analysts, with specific recommendations derived from proprietary analysis focused on individual companies. Managers consistently self-describe as having high-conviction, which is also verifiable on the basis that they only hold around 2/3 of the number of stocks as the average US-based mutual fund. This implies a set of highly active portfolio managers, selected for specific markets by institutional clients, who are meant to hold concentrated portfolios that deviate substantially from their benchmarks.

Unlike the managers studies in the majority of the mutual fund literature, our sample is not only focused on US equities. Many managers focus on their own domestic markets, including mandates focused on Europe, UK-specific, Australia, South Africa, and Middle Eastern markets, in addition to the US. Figure A.1 presents a breakdown of the average portfolio geographic exposure by year.

Our portfolios are also different from other databases of separately managed accounts used in the prior literature. For example, the managers in our sample have

Figure A.1. Breakdown of portfolio holdings by country over time

This figure presents a holdings-level breakdown of the sample over time. Much of the variation in the sample is driven by the fact that firms enter and exit the data over time, and so (for example) the Japan-focused managers are diluted as more firms enter the sample. The sample is based on the Worldscope classification of a security's domicile, and aggregated at the fund level as a percent of daily AUM, before being averaged by year across managers.



lower annualized portfolio turnover turnover–while the median annualized portfolio turnover is 35% in the Inalytics sample, it is 59% in the Morningstar Principia database (Chen et al. 2017). Managers in our sample also tend to be more concentrated, holding fewer names at a time. The average manager in the Inalytics database holds 78 distinct stocks, versus 103 stocks for managers in the Principia database (Chen et al. 2017). The portfolios in our sample also have lower average loadings on traditional style factors like value, size, and momentum, perhaps reflecting that these factors may be country specific rather than global in nature (Griffin 2002), and that our sample contains mandates which focus on specific regions other than the United States.

Next, we discuss some of the summary statistics which we compute from our position-level data and report in Panel B of Table A.1. Our simplest position-level variable is an indicator variable which equals 1 if the manager buys or sells a given

Table A.1. Summary statistics

This table reports summary statistics of the analysis dataset for 783 portfolios at various levels of aggregation. The position level summary statistics include various holding lengths, portfolio weights, future return measures and the number of trades (indicator for buy and sell trades). Future returns are reported in percentage points over specified horizons. The fund-level and position-level summary statistics are reported at monthly and daily frequencies, respectively. See Table A.2 and text for additional details on variable construction.

Variable	Count	Mean	Std	25th	50th	75th
Panel A: Fund level Summary						
Heuristics Intensity	0.5M	0.41	0.23	0.29	0.40	0.53
Weekly Fund Flows (%)	1.1M	0.35	5.2	-1.04	0.21	1.79
Concentration (%)	774	2.60	2.59	1.3	2.04	3.02
Large Block Holdings (%)	774	1.09	4.04	0.00	0.00	0.20
Average Size Quintile	783	4.42	0.63	4.23	4.63	4.89
Average Momentum Quintile	783	3.19	0.32	2.97	3.19	3.39
Average Value Quintile	783	2.72	0.50	2.42	2.77	3.02
Panel B : Position Level Summary						
Buying indicator	89.8M	0.03	0.16	0	0	0
Selling indicator	89.8M	0.02	0.15	0	0	0
Holding length since position open (days)	89.8M	484.4	512.9	119	314	679
Holding length since last trade (days)	89.8M	73.36	113.5	10	32	88
Holding length since last buy (days)	89.8M	112.3	152.4	18	57	144
Portfolio weight(%)	89.7M	1.20	1.61	0.24	0.79	1.65
1-day return (%)	82.1M	0.05	4.15	-1.11	0.01	1.17
Future 7-day return (%)	82.9M	0.21	5.83	-2.45	0.18	2.83
Future 28-day return (%)	82.8M	0.78	11.04	-4.63	0.81	6.18
Future 90-day return (%)	82.6M	2.56	20.16	-7.71	2.31	12.30
Future 180-day return (%)	81.5M	5.32	30.51	-10.46	4.16	18.88
Future 270-day return (%)	80.3M	7.87	38.54	-13.10	5.56	24.47
Future 365-day return (%)	78.9M	10.37	44.84	-15.08	7.24	29.73
Earnings announcement day indicator	49.3M	0.01	0.08	0	0	0
Active share	89.8M	0.86	1.27	0.11	0.55	1.28

stock on a given date. Of the 89 million position-date combinations in our sample where a stock was in the portfolio at either the start or end of the day, about 2.4 million of them involved an active purchase decision on that same day and 2 million of them involved active sell decisions, or about 2.6 percent and 2.2 percent of the time, respectively.

We compute several measures at the position level. First, we construct several

Table A.2. Summary of characteristics

This table describes how we construct several characteristics for use in our analysis.

Characteristics	Sorting	Construction
Cumulative Returns capped at K-days	Within Fund-date across stocks	$r^{cum}_{s,f,t} = \Pi^{i=t}_{i=t-\min\{K,d\}}(1 + r_{s,f,t}) - 1,$ where d is the time since a position opens.
Position past k day returns	Within Fund-date across stocks	$r^{past\ k}_{s,f,t} = \Pi^{i=t-1}_{i=t-k}(1 + r_{s,f,t}) - 1.$
Fund past k day returns	Across funds on daily basis	$r^{k}_{f,t} = \Pi^{i=t-1}_{i=t=k-1-1}(1 + r_{,f,t}) - 1.$
Heuristics Intensity	Across/Within funds on weekly/monthly basis	$\frac{\text{Total \# of Positions sold in Bin 1 or Bin 6 of past returns}}{\text{Total \# of Positions Sold}}.$
Position Size	Within Fund-date across stocks	$PositionSize_{s,f,t} = \frac{Quantity^{beginiing\ t}_{s,f,t} \times P_{s,f,t}}{Fund\ AUM_{s,f,t}}.$
Active share	Within Fund-date across stocks	Position size - weight in client-designated benchmark.
Net Buy	Within funds on weekly basis	# of stocks bought - # of stocks sold.
Monthly Turnover	Across funds on monthly basis	$turnover_{f,m} = \frac{\min\{total\ MarketValue^{buy}_{f,m} total\ MarketVvalue^{sell}_{f,m}\}}{\overline{MarketValue_{f,m}}}.$
Holding length last buy	Within Fund-date across stocks	# of trading days from last day on which a position was bought
Fund Flows	Within Fund across days	$flow_{f,t,k} = \frac{AUM_{f,t} - AUM_{f,t-k} r^{k}_{f,t}}{AUM_{f,t-k}}$
Position Size	Within Fund within date	$\frac{Shares_{f,i,t} * Price_{i,t}}{AUM_{f,t}}$
Trade Size	Within Fund within date	$\frac{\lvert Shares_{f,i,t} * Price_{i,t} - Shares_{f,i,t-1} * Price_{i,t} \rvert}{AUM_{f,t}}$
Value Quintile	By security within year	Quintile of Book to Market ratio relative to other securities in Worldscope
Size Quintile	By security within year	Quintile market capitalization relative to other securities in Worldscope
Momentum Quintile	By security within year	Quintile lagged annual return relative to other securities in Worldscope
Fund Characteristic Quintile	By Fund within date	Weighted average quintile of their holdings for the characteristic of interest (Value, Size, Momentum)
% Large Block Holdings	Calculated by fund/date averaged by fund	$\frac{\Sigma w_{i,f,t} I(w_{i,f,t} > 0.05)}{\Sigma w_{i,f,t}}$
% Long Term Holdings	Calculated by fund/date averaged by fund	% of fund AUM held at least 2 years
Concentration	Calculated by fund/date averaged by fund	Average weight of a security in the portfolio
Amihud Liquidity Measure	Calculated by security for each date	$\frac{DollarVolume}{\lvert Past30DayReturn \rvert}$

different measures of the holding length associated with a given position. Specifically, we consider the length of time (in calendar days) elapsed since the position was first added to the portfolio. In many case, this measure will be censored because a stock may have been in the portfolio since it was first added to our sample. The average holding length is 485 calendar days (or about 15 months), though this measure is downward-biased. As such, we also examine holding length measures which consider the time elapsed since a stock was most recently bought (or traded). The average position was last purchased about 112 calendar days (a bit less than four months) ago and was last traded about 10 weeks ago. In much of the analysis that follows, we will exclude stocks which were very recently bought to avoid having our results being driven by predictable buying (and lack of selling) behavior as managers split trades over several days while building up positions over time. Second, we compute the portfolio weight as a fraction of market value associated with each position on each date. The average stock has a weight of about 1.2 percent with a standard deviation of 1.6 percent. Inalytics also provides us with a measure of active share.

We also compute a number of backward or forward-looking return measures at the position level over various horizons, both overall and relative to the benchmark return. With the exception of 1-day measures (which refer to the prior trading day), we calculate horizons in calendar days.[58] For brevity, we only report summary statistics for forward-looking returns that are not adjusted for the benchmark. Volatilities of individual stocks are quite large, with a standard deviation of 45 percent at a 1 year horizon. As we discuss further below, we also consider several measures of prior position performance that are computed using periods of time which depend on holding period length.

There are several characteristics we calculate at the position level in order to explore differences between the investment approaches of different managers. We calculate the percentage of holdings which are long-term (held more than 2 years), and which are held in large blocks ($\geq 5\%$ of the portfolio AUM), and portfolio concentration (average weight of a security in a portfolio). Following the approach of Daniel et al. (1997), we also compute the value, size, and momentum quintiles for securities based relative to the global sample by looking each year where the security's characteristics fall in rank

[58]This choice is, in part, motivated by the fact that trading calendars differ slightly across exchanges. We take a number of precautions to reduce the potential influence of measurement errors in prices, including winsorizing 0.1 percent of returns in either tail by date. These steps are discussed at greater length in the Appendix A.1.

order relative to the rest of the Worldscope database (results are insensitive to computing these breakpoints at the nation, region, or aggregate level). For the purposes of making DGTW adjustments and/or forming counterfactual portfolios matched on characteristics, these security quintiles are computed annually using Worldscope characteristics (see Appendix A.3). For purposes of comparing across managers in the heterogeneity analysis in Figure 2, we simply average these assigned characteristics at the portfolio level across all positions held.

A.3 Adjusting performance measures for risk and other systemic factors

In this section, we provide additional details about how we construct counterfactuals to form "factor-neutral" portfolios. Specifically, we estimate stock-level exposures to the Fama-French/Carhart 4 factors using data from prior to the trade, then use these estimates to adjust our long short portfolios for ex-ante differences in these exposures.[59] For each stock-date, we subtract off the inner product of factor loadings and factor realizations, so

$$R_{i,t}^{FN} \equiv R_{i,t} - \Lambda'_{i,q(t)-1} F_t,$$

where $R_{i,t}$ is stock i's excess return on date t and F_t is a (4×1) vector of factor realizations.

$\Lambda_{i,q(t)-1}$ is a (4×1) vector of factor loadings which are estimated 1 year of daily data using data up to the end of the previous calendar quarter with Dimson (1979) adjustments for asynchronous trading. This adjustment includes one lead and lag for each factor, summing over the contemporaneous, lead, and lag betas. $R_{i,t}^{FN}$ captures return of a self-financing portfolio which, if factor loadings are estimated correctly and are stable, has zero exposure to the priced risk factors on each date. Thus, if the asset pricing model holds, all $R_{i,t}$ should earn zero excess return in expectation, and, accordingly, randomly sold portfolios should have the same factor-neutral returns period-by-period as actual stocks sold. Next, we compute value-added as before, by compounding factor neutral returns then compare cumulative factor-neutral returns of stocks traded with the average of cumulative factor-neutral returns of stocks held.

In order to test whether exposures to known characteristics are driving our results,

[59] The four factors are the market excess return, the Fama-French (2012) international size and value factors, as well as the Carhart momentum factor. As above, we compute loadings using data for the global factors from Ken French's website.

rather than only correcting for factor exposures, we also use the Worldscope data to construct characteristic-matched portfolios. To do this, we take the annual panel of characteristics for the securities, and construct 5 equal buckets along Book-to-Market (value), Market Capitalization (size), and the 1 year return for the previous year (momentum). This follows the methodology of (Daniel et al. 1997). We calculate portfolio returns for each portfolio in the 5x5x5 sort, and merge this data in to our panel of security returns (matching on the same quintile of characteristics). This allows us to calculate the excess returns of a security over the matched characteristic portfolio on a daily basis (referred to in the text as DGTW-adjusted returns or characteristic selectivity). The DGTW adjustment calculates the excess period return of the security over the matched portfolio; in our case we compute the excess return on a daily basis. We then take the excess returns for each security and aggregate with the same methodology as our other counterfactuals, grouping into equal-weight portfolios based on whether security was bought, sold, or neither on a given day.

A.4 Standard error computation for counterfactual return measures

In sections 3.2, 4, and 5, we estimate a number of average performance measures which are computed as a weighted average of cumulative long-short portfolio returns ($R_{buy} - R_{hold}$ and $R_{hold} - R_{sell}$). In this section, we describe our methodology for estimating double-clustered heteroskedasticity and autocorrelation robust standard errors associated with these mean calculations. We use a clustered standard error estimator which allows for autocorrelation and heteroskedasticity at the fund level following a similar approach to autocorrelation as Hansen and Hodrick (1980), who propose a standard error estimator for a general moving average process with a finite number of lags.

For each fund, let $\tau \in \{1, \dots, T_i\}$ index dates for which we have return measures for fund i, and $w_{i,\tau}$ be the weight associated with the τ^{th} observation for fund i and $t_i(\tau)$ be the mapping from the index τ to calendar time t for fund i. Our weighted mean calculation takes the following form for various choices of horizon H:

$$\bar{R} = \sum_{i=1}^{N} \sum_{\tau=1}^{T_i} w_{i,\tau} R_{i,t_i(\tau):t_i(\tau)+H}, \tag{1}$$

where $\sum_{i=1}^{N} \sum_{\tau=1}^{T_i} w_{i,\tau} = 1$ and $R_{i,t_i(\tau):t_i(\tau)+H}$ is the cumulative return measure for the associated portfolio for fund i from $t_i(\tau)$ to $t_i(\tau)+H$. Note that (1) also corresponds to the OLS formula for the following regression:

$$\sqrt{w_{i,\tau}}\ R_{i,t_i(\tau):t_i(\tau)+H} = \sqrt{w_{i,\tau}}\ \left(\bar{R} + \varepsilon_{i,t_i(\tau):t_i(\tau)+H}\right). \tag{2}$$

To implement a clustered estimator which accounts for both autocorrelation and cross-fund correlation, we also need to estimate

$$Cov[\varepsilon_{i,t(\tau_1):t(\tau_1)+H} \cdot \varepsilon_{i,t(\tau_2):t(\tau_2)+H}] = Corr[\varepsilon_{i,t(\tau_1):t(\tau_1)+H} \cdot \varepsilon_{i,t(\tau_2):t(\tau_2)+H}] \cdot Var[\varepsilon_{i,t:t+H}] \tag{3}$$

and

$$Cov[\varepsilon_{i,t(\tau_1):t(\tau_1)+H} \cdot \varepsilon_{j,t(\tau_2):t(\tau_2)+H}] = Corr[\varepsilon_{i,t(\tau_1):t(\tau_1)+H} \cdot \varepsilon_{j,t(\tau_2):t(\tau_2)+H}] \cdot \sqrt{Var[\varepsilon_{i,t:t+H}]Var[\varepsilon_{j,t:t+H}]}, \tag{4}$$

for cases when $i \neq j$. We assume that $\varepsilon_{i,t_i(\tau):t_i(\tau)+H}$ has mean zero in the cross-section and in the time series, but $Cov[\varepsilon_{i,t_i(\tau_1):t_i(\tau_1)+H}, \varepsilon_{i,t_i(\tau_2):t_i(\tau_2)+H}]$ may be larger than zero for $|\tau_1 - \tau_2| < H$ due to the overlapping structure in calendar time, and cross-portfolio correlations may be non-zero at each point in the time series for lags $< H$. Then, a general formula for the clustered standard error for equation (2) in this case (see, e.g., Cameron and Miller 2015) is

$$V_{clu}[\bar{r}] = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{\tau_1=1}^{T_i}\sum_{\tau_2=1}^{T_j} w_{i,\tau_1} w_{j,\tau_2} E[\varepsilon_{i,t(\tau_1):t(\tau_1)+H} \cdot \varepsilon_{j,t(\tau_2):t(\tau_2)+H}]\mathbf{1}[t(\tau_1), t(\tau_2) \text{ in same cluster}]}{\left[\sum_{i=1}^{N}\sum_{\tau_1=1}^{T_i} w_{i,\tau_1}\right]^2}. \tag{5}$$

Because of the highly unbalanced panel structure of the daily counterfactual returns data, many fund-specific autocorrelation estimates are informed by a very small number of observations. To improve the precision of our estimates, we pool information across funds to obtain more accurate estimates of autocorrelations while still allowing for heteroskedasticity. Exploiting the identity

$$Cov[\varepsilon_{i,t_i(\tau_1):t_i(\tau_1)+H} \cdot \varepsilon_{i,t_i(\tau_2):t_i(\tau_2)+H}] = Corr[\varepsilon_{i,t_i(\tau_1):t_i(\tau_1)+H} \cdot \varepsilon_{i,t_i(\tau_2):t_i(\tau_2)+H}] \cdot Var[\varepsilon_{i,t:t+H}], \tag{6}$$

we apply a common correlation term across funds but allow variances to vary at the fund level. We estimate a different autocorrelation for each lag length, which is measured as the number of weekdays between any two calendar time dates. We do this in three stages. First, at the fund level we estimate $Var[\varepsilon_{i,t_i(\tau_1):t_i(\tau_1)+H}]$. Next, on a fund-by-fund basis we estimate $Corr[\varepsilon_{i,t_i(\tau_1):t_i(\tau_1)+H} \cdot \varepsilon_{i,t(\tau_2):t(\tau_2)+H}]$ by taking the ratio of the sample mean of the cross terms at a given lag length to our fund-specific variance estimate.

We also need to account for cross-fund correlations, but because calculating it directly requires computationally infeasible combinatorics, we use a leave-one-out calculation, estimating the cross-correlation between a fund's period return and the period/lag returns of the weighted average of the other fund's portfolios. In other words, fixing the lag, we estimate

$$Corr\left[\varepsilon_{i,t(\tau_1):t(\tau_1)+H}, \left(\sum_{j\neq i} \frac{w_{j,t(\tau_2)}}{\sum_{k\neq i} w_{k,t(\tau_2)}} \varepsilon_{j,t(\tau_2):t(\tau_2)+H}\right)\right] \equiv Corr\left[\varepsilon_{i,t(\tau_1):t(\tau_1)+H}, \varepsilon_{\setminus i,t(\tau_2):t(\tau_2)+H}\right] \tag{7}$$

for each portfolio i. Then, finally, we aggregate to a single correlation number by taking a weighted average of the individual fund-level autocorrelation and cross correlation estimates and compute the standard error using equations (5-6) with the common correlation and fund-specific variance estimates, respectively.

Inspecting our autocorrelation estimates, it appears that the primary source of autocorrelation in our counterfactual performance measures comes from repeated trading of the same asset. For instance, if a PM sells a stock today, she is also somewhat more likely to sell it tomorrow and in the near future. Since we use cumulative return measures, consecutive trades of the same asset will mechanically be autocorrelated due to a moving average structure in the error term which comes from overlapping calendar time periods for cumulative returns.

Autocorrelation in trading appears to a very high frequency phenomenon. As a result, our estimated autocorrelation for our 1 year counterfactual sell measures is about 50% at a 1 trading day frequency and falls to 23% at a 5 trading day horizon. Within 1 and 3 months, these autocorrelations fall to about 7% and 3%, respectively, and decay rapidly towards zero thereafter. At long horizons, we find little to no evidence of autocorrelation even for periods which overlap in calendar time, as might be expected if stocks bought/sold tend to have different exposures to systematic risk relative to stocks not sold. In light of these estimates, since the computational cost of computing a full set of daily autocorrelations is quite high, we impose an estimate of zero after

100 business days, though overall estimates were quite similar when we allowed all autocorrelations for overlapping calendar time period to be nonzero.

Whereas these autocorrelation adjustments result in a nontrivial increase in estimated standard errors, the adjustment for cross-fund correlations turns out to be quite small. Since our PMs tend to hold concentrated portfolios that depart quite a bit from the benchmark, there is a fairly low (though not zero) overlap between holdings of different portfolios in our sample, and we do not find highly correlated trades in the cross-section. For factor-neutral portfolios, the average cross-fund correlation for the Sell - Hold portfolio when only looking at the same day of trading is 0.01. Over longer horizons, this estimate approaches closer to zero and even flips signs.

The small cross-fund correlations make sense considering the context of the sample—a set of concentrated managers with distinct mandates that are hired to earn excess returns by departing from the benchmark. For example, we observe roughly 18,000 different SEDOLs and 14,000 different ISINs held by the managers in our sample, though each only holds roughly 60 securities at a time. Cross covariances are essentially zero in magnitude, and, if anything, tend to be slightly negative after accounting for the correlation with lagged trades, with an average between-fund correlation of -0.005. In turn, the double clustered standard errors are quite similar and sometimes a bit smaller relative to the single clustered, HAC standard errors, as shown in Table A.3. Note that these results also speak against an explanation based on price impact from coordinated sales across funds. Given the small cross-fund correlations, PMs tend to not sell the same assets at the same time, implying that the price impact channel is unlikely to rationalize our performance results.

As an additional piece of evidence consistent with this argument, Appendix A.10 considers the returns from a set of strategies in which we form calendar time portfolios based on monthly signals which are constructed using our trading data. These portfolios earn quite high Sharpe ratios (though these calculations do not take into account transaction costs associated with implementing such a strategy), in large part because there are substantial benefits from diversifying across different stocks which are sold at given points in time. In addition, these portfolios have very low exposures to systematic factors, consistent with the arguments made throughout the paper that our sample of PMs mostly seek to generate superior performance via selecting individual stocks rather than by betting on systematic risk factors.

Table A.3. Standard Errors single-clustered HAC adjusted vs. double-clustered

Comparison of single-clustered HAC standard errors and double-clustered standard errors for the baseline results presented in Table 3. Double-clustered standard errors are extremely similar to single clustered, and sometimes are smaller due to negative between-fund correlations, as discussed in Appendix A.4.

	Panel A: Buy Standard Errors		**Panel B: Sell Standard Errors**	
Factor-Neutral				
Window	Single-Clustered	Double-Clustered	Single-Clustered	Double-Clustered
28	0.04	0.04	0.04	0.04
90	0.08	0.09	0.08	0.09
180	0.13	0.15	0.14	0.15
270	0.17	0.20	0.18	0.18
365	0.22	0.24	0.22	0.20
Unadjusted				
Window	Single-Clustered	Double-Clustered	Single-Clustered	Double-Clustered
28	0.04	0.05	0.04	0.04
90	0.08	0.10	0.08	0.10
180	0.13	0.18	0.14	0.17
270	0.17	0.22	0.18	0.20
365	0.22	0.27	0.22	0.23

A.5 Additional details and robustness checks for performance results

Due to concerns that trading frictions may prevent managers from exiting or entering positions as they see fit (and thus a random security may not be a good counterfactual), we merge in both Worldscope and Datastream to the sample to drop securities which do not meet a minimum threshold for either market capitalization or liquidity. In both cases, we take a measure (either market capitalization or the Amihud liquidity measure), calculate a 5th percentile breakpoint for the universe of securities in either Worldscope (for market cap) or Datastream (for Amihud liquidity) for each year, and merge the relevant sample into our portfolio holdings. We perform tests where we exclude all securities below a certain threshold from all relevant portfolios (buy, sell, and hold) and where we exclude securities only from the hold portfolio (which would be the relevant counterfactual). For one result in the appendix, we also merge NYSE breakpoints taken from Ken French's data library, using the 5th percentile breakpoint from those data.

Because a random counterfactual may not be appropriate (for example, if a portfolio

Table A.4. Post-trade unadjusted returns relative to counterfactual, additional robustness checks

This table presents the average value added measures (post-trade unadjusted returns relative to a random sell counterfactual) for buy and sell trades. We show all unadjusted versions of the specifications in Table 3 which were only presented in factor neutral returns.

Performance	Panel A: Buy					Panel B: Sell				
Measure	28	90	180	270	365	28	90	180	270	365
Drop 5% smallest securities	0.19	0.41	0.46	0.56	0.60	-0.05	-0.23	-0.76	-0.75	-0.85
	(0.05)	(0.10)	(0.17)	(0.21)	(0.25)	(0.04)	(0.10)	(0.15)	(0.19)	(0.24)
Drop 5% most illiquid securities	0.41	0.70	0.92	1.08	0.26	0.03	-0.15	-0.37	-0.42	-0.60
	(0.05)	(0.11)	(0.19)	(0.24)	(0.31)	(0.04)	(0.09)	(0.16)	(0.20)	(0.24)
Value-weighted	0.36	0.75	1.06	1.32	1.56	-0.01	-0.27	-0.62	-0.85	-1.00
	(0.05)	(0.15)	(0.30)	(0.39)	(0.47)	(0.05)	(0.13)	(0.26)	(0.34)	(0.39)
Drop most illiquid quintile by portfolio	0.35	0.61	0.93	1.14	1.33	0.05	-0.17	-0.35	-0.36	-0.50
	(0.05)	(0.10)	(0.17)	(0.23)	(0.30)	(0.04)	(0.09)	(0.14)	(0.18)	(0.22)
Drop shortest quintile of holding length	0.41	0.77	1.07	1.27	1.49	0.03	-0.18	-0.55	-0.70	-0.81
	(0.04)	(0.10)	(0.18)	(0.24)	(0.28)	(0.04)	(0.10)	(0.17)	(0.20)	(0.23)
Drop week post trade, Unadjusted Returns	0.21	0.52	0.74	0.88	1.05	-0.03	-0.21	-0.53	-0.71	-0.82
	(0.04)	(0.09)	(0.16)	(0.21)	(0.25)	(0.04)	(0.09)	(0.17)	(0.21)	(0.24)
Match counterfactual on										
Market capitalization quintile	0.33	0.50	0.54	0.61	0.67	-0.16	-0.41	-0.44	-0.51	-0.62
	(0.05)	(0.10)	(0.18)	(0.23)	(0.28)	(0.04)	(0.09)	(0.14)	(0.18)	(0.22)
Liquidity quintile	0.38	0.65	0.86	1.00	1.19	0.05	-0.18	-0.46	-0.49	-0.63
	(0.05)	(0.11)	(0.19)	(0.25)	(0.31)	(0.05)	(0.10)	(0.18)	(0.22)	(0.25)
Idiosyncratic volatility quintile	0.39	0.68	0.90	1.01	1.02	0.00	-0.23	-0.56	-0.76	-0.92
	(0.04)	(0.09)	(0.14)	(0.19)	(0.22)	(0.04)	(0.09)	(0.15)	(0.18)	(0.21)
Position size quintile	0.32	0.51	0.57	0.71	0.94	0.03	-0.17	-0.56	-0.79	-0.92
	(0.05)	(0.09)	(0.13)	(0.18)	(0.24)	(0.04)	(0.09)	(0.15)	(0.17)	(0.20)
Previous quarter's benchmark adjusted return	0.38	0.73	0.99	1.23	1.45	0.07	-0.08	-0.33	-0.49	-0.66
	(0.04)	(0.09)	(0.13)	(0.17)	(0.23)	(0.04)	(0.08)	(0.13)	(0.17)	(0.21)

Table A.5. Post-trade factor-neutral returns relative to counterfactual, additional robustness checks

This table presents the average value added measures (post-trade returns relative to a random sell counterfactual) for buy and sell trades. We show several additional robustness checks on a factor-neutral basis.

Performance	Panel A: Buy					Panel B: Sell				
Measure	28	90	180	270	365	28	90	180	270	365
Drop 5% smallest, NYSE breakpoints	0.16	0.33	0.25	0.44	0.48	0.01	-0.12	-0.55	-0.49	-0.69
	(0.05)	(0.10)	(0.17)	(0.21)	(0.26)	(0.04)	(0.10)	(0.16)	(0.20)	(0.23)
Drop 20% smallest, NYSE breakpoints	0.28	0.49	0.59	0.61	0.58	0.00	-0.24	-0.55	-0.62	-0.76
	(0.04)	(0.09)	(0.15)	(0.19)	(0.25)	(0.02)	(0.09)	(0.12)	(0.17)	(0.20)
Drop most illiquid quintile by portfolio	0.41	0.70	0.92	1.08	1.26	0.03	-0.15	-0.38	-0.42	-0.60
	(0.05)	(0.11)	(0.19)	(0.24)	(0.31)	(0.04)	(0.09)	(0.16)	(0.20)	(0.24)
Drop 5% most illiquid from counterfactual	0.35	0.61	0.89	1.05	1.29	0.02	-0.24	-0.49	-0.55	-0.71
	(0.05)	(0.09)	(0.15)	(0.20)	(0.26)	(0.05)	(0.10)	(0.16)	(0.20)	(0.22)
Drop 5% smallest from counterfactual	0.34	0.57	0.73	0.87	0.91	0.00	-0.30	-0.69	-0.88	-1.03
	(0.04)	(0.09)	(0.16)	(0.21)	(0.26)	(0.04)	(0.11)	(0.16)	(0.20)	(0.23)
Omit largest holdings quintile	0.31	0.58	0.87	1.11	1.37	-0.05	-0.42	-0.75	-0.98	-1.17
	(0.06)	(0.13)	(0.27)	(0.38)	(0.46)	(0.06)	(0.18)	(0.37)	(0.44)	(0.45)
Match counterfactual on										
Market capitalization quintile	0.28	0.37	0.44	0.50	0.51	0.00	-0.22	-0.47	-0.48	-0.58
	(0.04)	(0.09)	(0.13)	(0.16)	(0.19)	(0.05)	(0.09)	(0.16)	(0.20)	(0.26)
Book to market quintile	0.35	0.56	0.66	0.78	0.92	0.03	-0.13	-0.49	-0.45	-0.59
	(0.05)	(0.09)	(0.14)	(0.19)	(0.23)	(0.04)	(0.10)	(0.15)	(0.18)	(0.22)
Momentum quintile	0.32	0.49	0.56	0.62	0.73	0.01	-0.20	-0.54	-0.66	-0.68
	(0.05)	(0.09)	(0.15)	(0.19)	(0.24)	(0.04)	(0.09)	(0.15)	(0.18)	(0.20)
Liquidity quintile	0.32	0.53	0.82	0.98	1.19	0.04	-0.22	-0.43	-0.48	-0.56
	(0.05)	(0.10)	(0.17)	(0.22)	(0.28)	(0.05)	(0.10)	(0.16)	(0.20)	(0.23)

was only comparing securities with similar market capitalizations) we perform a series of matching exercises, where we only compare securities to similar securities matched on characteristics. In order to construct those counterfactuals, we construct quintiles for each portfolio-date along some characteristic of interest, and use the portion of the portfolio which is held (and not traded) within that same quintile as a counterfactual. When aggregating across these quintiles, we weight the average by the number of buy or sell trading days within a specific quintile. Quintiles are calculated based on the source data, grouping annually. For example, Worldscope provides an annual panel of characteristics, and we compute quintiles by year before merging with our sample.

Despite the relatively small AUMs (compared to large mutual funds), one may be concerned that funds may have the potential for price impact when buying or selling. To address this concern, we have constructed a number of counterfactuals that start a certain number of days after the trade. In the paper we report several of these specifications, which are all constructed in the same way other than the number of days omitted. While we only report results for one week for brevity, omitting the first few days post-trade has the tendency to make selling performance look worse, not better, which is the opposite of what we would predict from a price impact mechanism.

This appendix also reports several additional robustness exercises that address price impacts and illiquidity. First, for each of the specifications in Table 3 that only reported results with factor neutral returns, A.4 reports the same specifications with unadjusted returns. In A.4 and A.5 we also perform a wide variety of additional tests. The first batch are focused on issues of liquidity, e.g. that managers are unable to sell securities from the counterfactual due to liquidity concerns, or that there may be large price impacts in illiquid markets. In addition to the unadjusted version of the tests reported in Table 3, we conduct several additional tests. First, we use NYSE 5% breakpoints (rather than the Worldscope 5% breakpoints), which have a higher marketcap threshold due to the fact that NYSE stocks are larger than most of the emerging markets equities in our sample. We also include more aggressive versions of the analysis which drop the 20% least liquid portion of a given portfolio's holdings on a given day, finding similar results as our main specification.

In addition, because the portfolio managers *do* choose to trade the securities in the buy and sell portfolios, the liquidity concerns for the unobserved trades must be in the set of holdings that were not traded on a given day. It is plausible that managers have skill when choosing to trade illiquid securities, or are intentionally providing liquidity to the market at a specific time to generate excess returns. Thus, we also perform tests where we only exclude small and illiquid securities from the counterfactual portfolio, finding that the results are similar to the original specification, often with larger magnitudes than the baseline result.

It may also be the case that portfolio managers have specific style tilts, tight tracking error budgets, or position size limits. We perform additional tests to make sure that those potential issues are not driving our result. Rather than assigning a random counterfactual, we condition on the counterfactual matching the quintile of the sold security along the respective characteristics, including position size, market capitalization, book to market ratio, momentum, and liquidity. We cannot jointly test these because the dimensionality of a tri-variate sort quickly exceeds the average number of securities held by a given fund, but each univeriate matched counterfactual has similar results to the baseline. The fact that, after accounting for factors at the security level *and* matching the counterfactual on traditional style characteristics, we still observe the same performance results for both buying and selling should also assuage concerns that any systematic return premium is driving our findings.

Finally, we directly address potential position size limits by excluding the largest quintile of holdings in each portfolio from the buy and sell portfolios. If sells that lacked proper attention were driven by automatic rebalancing based on upper limits for a given position, we would expect that underperformance to be driven by "forced trades" of a PM's largest positions, measured as a proportion of their AUM on the day prior to the trade. We find that, if anything, selling underperformance becomes more pronounced when we exclude the top 20% of a portfolio's securities. This implies that, consistent with our attention mechanism, sales of a PM's largest positions (which are likely to be attended to) tend to look better sells relative to the average sell. Accordingly, when we drop the largest quintile of position size from the analysis (trade and counterfactual portfolios), we see that average selling performance measures deteriorate even further.

We can also conduct a simple model-free diagnostic which simulaneously addresses two potential alternative explanations for our results mentioned in the main text. First, results could driven by common exposures of stocks traded to common systematic, priced factors which were omitted from our explicit risk-adjustment procedures. Second, tracking error constraints might imply that a stock sold might need to share common systematic risk exposures with a stock purchased (e.g., to hold constant exposure to a given industry or region within the portfolio), leading the stocks sold to potentially share exposures to these factors (which subsequently outperform) as stocks purchased. We can address both of these potential arguments to some extent without needing to take a stand on what these specific systematic factors might be. The basic logic is as follows: if stocks purchased and stocks sold shared a common exposure to an unobserved risk factor which was not also shared by stocks held, we would expect to see that stocks bought and sold contemporaneously in time to have returns that are fairly highly correlated with one another, even after adjusting for exposures to common risk factors that appear in our factor-neutral portfolios.

To get at this question, we compute the correlation between R_{buy} and $R_{sell} - R_{hold}$ for stocks that occur on the same day or within a few trading days. Intuitively, if the stocks sold tend to share common systematic risks with stocks bought, we would expect to see a large positive correlation between R_{buy} and R_{sell}—a much larger one than between R_{hold} and R_{sell}. To get an idea of some magnitudes that we might expect to see, (Campbell, Lettau, Malkiel, and Xu 2001) decompose stock returns into market, industry-specific, and firm-specific components, and find that the variance of the industry-specific component (after residualizing with respect to the market and using the fairly coarse Fama-French 49 industry classification) is around 0.001 at a monthly frequency. Likewise, the firm-specific component has a variance of around 0.005. This would imply that the R^2 of a regression of a firm stock return (after absorbing the market) on an industry-specific portfolio would be in the neighborhood of 0.17, or, in other words, the correlation between the firm-specific component and the industry-specific component is around 40%.

In our data, the contemporaneous correlation of $R_{buy} - R_{hold}$ and $R_{sell} - R_{hold}$ when buys and sells happen on the same day is between 1.5% and 2.5%, depending on the future return horizon we look at, implying an R^2 of at most about 0.0006. This suggests that the stocks bought have essentially the same covariance with stocks sold as stocks held (which could have been sold instead). These correlations are essentially zero for trades occurring one or more trading days apart from one another. Such magnitudes are far too small to support an explanation based on exposures to unobserved systematic factors.

The results thus far have examined performance of all buying and selling decisions together. However, both buys and sells differ in the extent to which they add or subtract from the portfolio. Some buys add a little bit to an existing position while others introduce a substantial amount of shares or start a whole new position in the

Table A.6. Post-trade returns relative to counterfactual for trades with $\geq$ %50 change in weights

This table presents the average value added measures for large trades ($\geq$50% change in the weight of a position) under two measures of returns 1) raw returns and 2) factor-neutral returns. Double-clustered standard errors, computed using the method described in the section A.4, are reported in parentheses.

	Panel A: Large Buys					Panel B: Large Sells				
	28	90	180	270	365	28	90	180	270	365
Unadjusted	0.60	1.04	1.28	1.45	1.72	-0.43	-0.67	-1.05	-1.25	-1.27
	(0.07)	(0.17)	(0.29)	(0.35)	(0.42)	(0.09)	(0.27)	(0.50)	(0.75)	(0.86)
Factor-Neutral	0.54	0.81	1.04	1.18	1.44	-0.45	-0.70	-1.09	-1.19	-1.11
	(0.05)	(0.09)	(0.19)	(0.25)	(0.31)	(0.08)	(0.22)	(0.48)	(0.65)	(0.70)

portfolio; similarly, some sells cut a bit from existing positions while others unload substantial shares or cut the asset altogether. We look separately at decisions which change the weight of a position by 50 percent or more, relative to the current weight of a position in the portfolio. For example, a sell which changes the relative weight of a position from 10% to 5% or a buy which changes a position's weights from 20% to 40% would fall into this category. We also include new purchases and full liquidations in this category. Table A.6 above presents these results relative to the same counterfactual used in Table 3. The same pattern of results is observed.[60]

A.6 Additional Institutional Context: Reporting of Buy vs Sell Performance

In this section, we discuss the common approaches to reporting performance and the extent to which they facilitate the identification of value-added from buying and selling decisions. PMs and their clients generally use Brinson-style attribution to report and monitor their portfolio performance. This method looks at sector/country/style characteristics and measures the portfolio's over- or under-weights relative to a benchmark. The Brinson-Fachler model for attribution, which is the industry standard method of performing this sort of analysis, breaks down the excess return of a portfolio over its benchmark into an "allocation" effect, a "selection" effect, and an "interaction" effect.

[60]Note that the directionally greater performance of large relative buys compared to other buy trades is consistent with PMs trading on information when opening or increasing a position by a significant amount.

The allocation effect is constructed by multiplying the extent to which a portfolio overweights a given sector (or other category such as country) by the extent to which the benchmark return for that sector outperformed the overall benchmark return. This is carried out separately for each sector, and may be added for all sectors to give an overall allocation effect. Having a separate allocation and selection effect separates decisions related to which sector the PM chose to invest in from the actual names picked within each sector.

The selection effect assumes the sector weights of the benchmark, but uses the aggregate portfolio returns in each sector in place of the benchmark returns, and then looks at the difference. This component is meant to identify poor security selection. However, because this uses aggregate returns, the only context in which it could identify selection underperformance is if the portfolio, on average, holds worse securities than the benchmark does, within any given sector. In other words, even if recent sell trades hurt performance by prematurely reducing exposure to a security that outperformed, these attribution measures would still give the manager credit for positive security selection because they are based on portfolio weights, not changes in weights.

As a consequence of this reporting standard, even if a portfolio is underperforming its benchmark after adjusting for industry weights, the emphasis of such attribution is the current holdings. The only way that costs from selling could be easily isolated is if the portfolio previously had a large position in a security that also had a large weight in the benchmark, and also that such a security was one of the largest contributors to the benchmark return. In such a case, we believe it to be plausible that a manager/client would take another look at their selling decision, especially because underweighting a security which has a large position in the benchmark is typically an active choice. However, because of the nature of the management style (concentrated portfolios that depart from benchmark), most of the securities that are sold are not large positions in the benchmark.

Given the emphasis on performance of existing holdings, this reporting standard makes it much more difficult to identify underperformance in sales than in buys, which naturally show up as current holdings. In turn, managers are less likely to notice that they are using costly heuristics when selling, which perpetuates underperformance relative to counterfactual measures. Noticing these costs is especially challenging in light of the horizon over which underperformance takes place. While we find that these potentially inattentive selling decisions have neutral performance over short horizons (e.g., less than 1 month), they substantially underperform over longer ones. This added delay between the decision and the realization of the outcome further adds to the difficulty of learning from prior decisions.

A.7 Additional results and information related to trading probabilities

This section provides additional details about our analysis in section 4.4.1. First, we discuss how the bins are constructed, then report additional results related to the probability of making buy and sell decisions, mainly conditional specifications which separate out large trades, sort on position size and holding length, or condition on a day being a "trading day."

On each trading date, we sort stocks into N_{bin} bins using these relative rankings. We always choose an even number of bins and always set the breakpoint between bins $N_{bin}/2$ and $N_{bin}/2+1$ equal to zero. This ensures that all stocks in bins $N_{bin}/2$ have declined relative to the benchmark. We choose all remaining breakpoints so that (ignoring issues related to discreteness) there are equal numbers of stocks in bins $1,\ldots,N_{bin}/2$ and bins $N_{bin}/2+1,\ldots,N_{bin}$. As a baseline, we consider $N_{bin}=20$. Some specifications collapse across bins to fit more conveniently in tabular format—the results are always robust to the number of bins (N_bin) considered.

Throughout out analysis of trading probabilities, we make one substantive restriction on the sample of stocks which are under consideration. In predicting the probability that a manager will add to/reduce an existing position, we exclude stocks that were bought in the very recent past. Specifically, we sort positions based on the holding length since the last buy trade and exclude the shortest 20% (shortest time elapsed since last purchase) from our calculations. We do this to avoid a fairly mechanical relationship between our prior return measure, which has a variance that shrinks with the holding period, and the probability of buying/selling that can be generated if managers build up positions by splitting buy trades over short windows of time.[61] Such trades likely originate from a single purchase decision being executed over time, and so we construct our measures to treat them as such. Further, to ensure meaningful distinctions between bins, we exclude fund-dates which include fewer than 40 stocks in the portfolio, though results for predicted selling probabilities do not meaningfully

[61] This phenomenon mechanically tends to increase the likelihood that positions with non-extreme returns are bought and decrease the likelihood that they are sold, since a manager is unlikely to sell an asset immediately after or while actively building a position in it.

Table A.7. Probability of buying and selling: larger consideration set and large trades

This table presents the probability of buying for security in the investors holdings and consideration sets, as well as the probability of trades with a $\geq$ %50 change in weights as a function of prior returns. Panel A presents reports differences in probabilities, in percentage points, of buying selling by bins of past benchmark-adjusted returns for stocks in the larger consideration set (current holdings plus stocks added to the portfolio in the next year) by 20 bins of positions past benchmark-adjusted returns. The baseline probability of trading a stock in the omitted category, bin 10, is reported in the rightmost column. The first row applies the same filters as Figure 2, while the second row considers a wider set of stocks following the approach described in the main text. Panel B presents the relative probability of large buys and sells by 20 bins of past positions benchmark-adjusted returns capped at 90 days.

Panel A: Buying probability by bins of prior 1 quarter benchmark-adjusted returns

	Differences relative to bin 10 for prior return bins								Baseline
Buying prob for	1	2	3-5	6-9	11-15	16-18	19	20	prob (bin 10)
Current holdings	0.03%	0.05%	0.03%	0.00%	0.02%	0.00%	-0.02%	-0.06%	1.10%
Consideration set	0.09%	0.07%	0.05%	0.00%	0.00%	0.07%	0.10%	0.06%	1.62%

Panel B: Trading probability by bins of prior benchmark-adjusted returns, capped 1 quarter

	Differences relative to bin 10 for prior return bins								Baseline
Trading Probability	1	2	3-5	6-9	11-15	16-18	19	20	prob (bin 10)
Large Buys	0.02%	0.02%	0.01%	0.02%	0.01%	0.01%	0.01%	0.01%	0.16%
Large Sells	0.38%	0.27%	0.21%	0.07%	0.01%	0.01%	0.03%	0.12%	0.37%

change without such a restriction.

In the main text, we focus on PMs' decisions to purchase/sell existing holdings. Since PMs can add new positions to their portfolios, it is highly likely that they have a wider consideration set than existing holdings which is relevant from their buying decisions. Here, we describe a complementary approach which also allows for new stocks to be added to the portfolio in addition to current holdings. This approach includes all purchase decisions—including the opening of brand new positions—and calculates relative prior returns by broadening the consideration set to assets that are likely being considered for purchase. For this wider consideration set approach, we use the prior 1 quarter benchmark-adjusted return. Specifically, because our dataset contains not only current and past holdings for each PM but future holdings as well, we can include assets that the PM is likely considering by looking at what he ended up buying within 12 months of the current date. We include those assets in the portfolio when computing the prior return bins to examine whether new positions are more or less likely to be bought depending on prior returns relative to the larger consideration set.

In Table A.7, Panel A replicates the buying decisions presented in Figure 3 but includes new buys using our second approach which expands the consideration set. Specifically, we report differences in probabilities relative to a baseline category (bin 10, stocks which barely underperformed the benchmark) of buying across categories of prior returns. For ease of comparison, the top row reports our estimate from Panel B of Figure 3, which uses the 1 quarter prior benchmark-adjusted return measure as the sorting variable. We average probabilities across several intermediate bins for brevity, and report the baseline probability associated with the omitted category in the final column. Then, the second row uses the same sorting variable but also includes stocks in the broader consideration set (as defined above) and eliminates our restriction which excluded stocks in the bottom bin of holding length since last buy. We see that the probabilities of purchasing an asset remain quite flat with respect to prior returns.

Panel B of Table A.7 depicts the propensity to engage in large transactions—selling more than 50 percent of an asset—as a function of prior returns. We see a similar U-shape emerge as when we consider the all sales together. Again, this pattern is not matched for the probability of engaging in large purchases, which remain quite flat across bins of prior returns. Together, these results demonstrate that we can predict selling decisions based on observables from the PM's current holdings with some confidence; in contrast, these observables—nor any others that we have considered—do not predict buying decisions.

We also examine whether our observed pattern can be explained by other variables which could be correlated with our prior return measures: holding length and position size. In the former case, as discussed above, positions which have only been held for a short period of time will to have less dispersion in returns and also may be more likely to be bought and less likely to be sold. In the latter, even if initial positions all begin at the same size, extreme returns could lead them away from target weights, and rebalancing motives could drive managers to sell positions with extreme positive returns that have become too large.[62]

[62]Note, however, that similar logic would potentially imply that we would see less selling of positions that have became small due to portfolio drift, which we do not observe. Also, from the univariate evidence above, we do not see large increases in buying for positions that declined in value, as would be predicted by this channel. In regressions below, we will always include linear and quadratic controls for position size and holding length.

Table A.8. Probability of trading by prior returns and position characteristics

This set of tables reports differences in probabilities, in percentage points, of buying/selling by bins of past benchmark-adjusted returns double sorted with bins of position characteristics holding length and position sizes relative to bin 10 of past benchmark-adjusted returns within each category. The top section of each panel reports relative probabilities of buying and the bottom section reports relative probabilities of selling. Baseline probabilities for the omitted category are reported below. Columns represents different holding lengths in Panel A and position sizes in Panel B. Different bins of past position returns are reported in rows, together with the baseline probability of the omitted category.

Panel A: Holding Length

Trade	Past Return\Holding Length	Shortest	Short	Short-Med	Med-Long	Long	Longest
	1	-2.34	-0.20	-0.15	-0.07	-0.03	0.03
	2	-1.99	-0.05	-0.03	0.02	0.05	0.07
	3-5	-2.08	-0.02	-0.01	0.05	0.04	0.05
	6-9	-1.20	0.05	0.05	0.13	0.07	0.04
	11-15	0.28	-0.03	-0.03	0.01	0.04	0.05
Buy	16-18	-1.59	-0.19	-0.12	-0.03	0.02	0.08
	19	-2.27	-0.33	-0.21	-0.06	0.02	0.09
	20	-2.93	-0.48	-0.29	-0.10	-0.02	0.09
	Baseline: 10	8.78	2.03	1.56	1.22	0.86	0.68
	1	0.45	0.78	1.17	1.14	1.52	1.70
	2	0.31	0.40	0.59	0.64	0.93	1.05
	3-5	0.24	0.24	0.35	0.34	0.42	0.58
	6-9	0.16	0.14	0.19	0.18	0.12	0.16
Sell	11-15	0.07	0.08	0.08	0.02	0.05	0.01
	16-18	0.32	0.27	0.32	0.32	0.35	0.34
	19	0.49	0.47	0.61	0.56	0.63	0.58
	20	0.75	0.94	1.12	1.15	1.25	1.21
	Baseline: 10	1.02	1.63	2.07	2.15	2.17	2.33

Panel B: Position Size

Trade	Past Return\Position Size	Smallest	Small	Small-Med	Med-Large	Large	Largest
	1	-0.21	-0.07	0.11	0.19	0.31	0.49
	2	-0.12	-0.01	0.11	0.15	0.25	0.39
	3-5	-0.11	-0.03	0.05	0.12	0.18	0.27
	6-9	0.05	0.02	0.10	0.06	0.07	0.13
	11-15	0.08	0.01	0.05	0.03	-0.02	-0.02
Buy	16-18	0.02	-0.05	-0.01	-0.02	-0.05	-0.08
	19	-0.07	-0.09	-0.04	-0.04	-0.07	-0.14
	20	-0.10	-0.13	-0.09	-0.09	-0.13	-0.20
	Baseline: 10	0.98	1.07	1.03	1.07	1.17	1.28
	1	1.33	0.90	0.93	1.02	1.02	0.90
	2	0.92	0.54	0.55	0.65	0.64	0.53
	3-5	0.55	0.30	0.31	0.35	0.37	0.30
	6-9	0.17	0.13	0.12	0.12	0.12	0.09
	11-15	-0.06	0.04	0.03	0.10	0.13	0.14
Sell	16-18	0.07	0.23	0.30	0.42	0.51	0.59
	19	0.22	0.41	0.51	0.71	0.83	0.96
	20	0.77	0.97	1.13	1.28	1.40	1.66
	Baseline: 10	3.38	1.93	1.81	1.85	1.95	2.18

As in Table 3, we assign each stock into one of 20 bins based on prior returns and the other sorting variable, respectively. Since the break points used for the second characteristic are the same regardless of the bin associated with the first characteristic, there will be unequal numbers of observations in each bin. We then report the buying (top panel) or selling (bottom panel) probabilities within each group relative to the middle, least extreme bin (bin 10). As in Figure A.7, we average across several intermediate categories and separately report the probability of trading for the omitted category.

Panel A of Table A.8 double sorts on six bins based on time elapsed since last buy (the variable we filter on) and prior returns. For this analysis only, we do not discard any stocks from the analysis based on the holding period measure. One can observe the mechanical patterns discussed in Section 4.4.1 when looking at the buying probabilities of assets in the bin with the shortest holding length; buying probabilities are quite flat in prior returns for longer holding periods. In contrast, assets in extreme bins are much more likely to be sold across all holding lengths.

As shown in Panel A of Table A.8, we observe that selling probabilities feature a pronounced U-shape for all position sizes, a pattern that holds robustly within all position size bins. For larger positions, we see some evidence of PMs adding to their biggest positions following losses. However, even within these categories, buying probabilities increase gradually with losses and decrease gradually with gains, whereas corresponding sell probabilities increase much more dramatically for the more extreme return categories. Additionally, the magnitudes for buys are generally much smaller compared to the respective selling probabilities. We discuss position size more in Section 4.1 of the main text.

Table A.9 presents results for the same specification as the linear probability models in Table 5, but conditions on there being at least one trade. In other words, this asks the question, "on days where a PM buys, how do the past extreme returns influence her decision." For sells, this instead conditions on sells. We find that results are quite similar to those in Table 5, but with larger magnitudes, especially for the selling results.

A.8 Additional results on the attention mechanism

In this section, we present additional evidence that the poor selling skill is driven by a lack of attention to those decisions. Namely we separate out periods where portfolio

Table A.9. Probability of trading based on prior returns, conditional on buy or sell

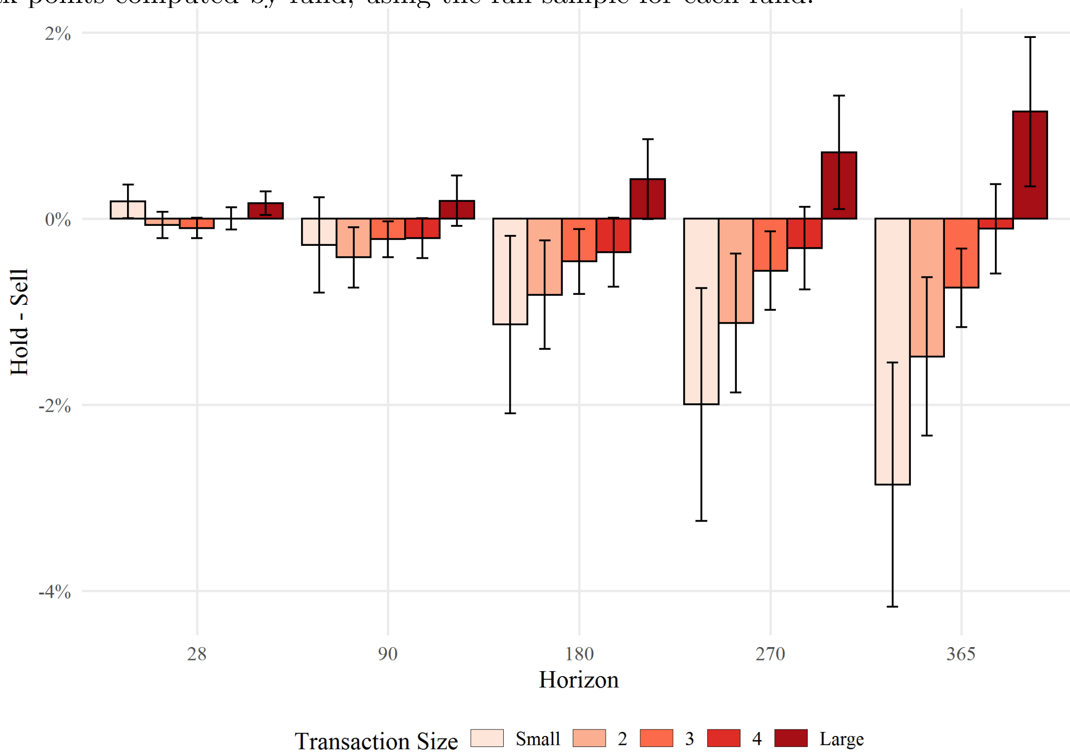
This table presents position-level estimates of a linear probability model (in percentage points) for the likelihood of trading a given stock. The key explanatory variables of interest are indicators corresponding to 20 bins of past benchmark-adjusted returns capped at one year, where the tenth bin is the omitted category. We control for fund characteristics including lagged log(AUM), prior-month turnover, the annual volatility of a funds benchmark-adjusted returns, and prior month loadings on Fama-French Cahart regressions (calculated using the Dimson (1979) procedure using 1 year of prior daily returns). We adjust for position-level characteristics including linear and quadratic terms in holding lengths (overall and since last buy) and position sizes(% AUM) at the beginning of the day. The coefficients and t-statistics are reported for the estimates on each bin.

	Buying Probability Given Buy Day				Selling Probability Given Sell Day			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bin 1	-0.058	-0.068	-0.108**	-0.282*	2.476***	2.455***	2.177***	1.382***
	(-1.574)	(-1.821)	(-2.845)	(-2.409)	(15.105)	(14.925)	(13.598)	(6.484)
Bin 2	0.067*	0.065	0.057	0.024	1.452***	1.438***	1.189***	0.881***
	(2.068)	(1.719)	(0.697)	(-1.843)	(12.849)	(12.627)	(10.777)	(5.803)
Bin 3 to 5	0.071**	0.063*	0.047	-0.041	0.761***	0.752***	0.597***	0.467***
	(3.005)	(2.567)	(1.852)	(-0.801)	(10.928)	(10.666)	(9.161)	(5.064)
Bin 6 to 9	0.033*	0.028	0.021	-0.009	0.170***	0.166***	0.128***	0.084
	(2.245)	(1.787)	(1.340)	(-0.328)	(5.833)	(5.477)	(4.769)	(1.934)
Bin 11 to 15	-0.073***	-0.078***	-0.057***	-0.034	-0.007	0.010	0.017	0.115**
	(-4.110)	(-4.143)	(-3.741)	(-1.166)	(-0.194)	(0.283)	(0.526)	(2.618)
Bin 16 to 18	-0.177***	-0.185***	-0.183***	-0.203**	0.488***	0.500***	0.424***	0.928***
	(-5.816)	(-4.668)	(-6.291)	(-3.679)	(6.794)	(6.831)	(6.240)	(8.674)
Bin 19	-0.240***	-0.248***	-0.255***	-0.264**	0.971***	0.978***	0.870***	1.448***
	(-6.678)	(-6.705)	(-7.497)	(-3.043)	(10.039)	(9.985)	(9.295)	(9.784)
Bin 20	-0.300***	-0.308***	-0.345***	-0.304**	2.081***	2.082***	1.916***	2.070***
	(-7.254)	(-7.271)	(-8.432)	(-2.679)	(15.377)	(15.245)	(14.446)	(10.492)
Fund Control	Yes	Yes	No	Yes	Yes	Yes	No	Yes
Fixed Effects	None	Date	Fund × Date	Stock × Date	None	Date	Fund × Date	Stock × Date
R^2	0.034	0.041	0.285	0.346	0.005	0.009	0.164	0.390
N	29.3M	29.3M	30.2M	22.4M	27.8M	27.8M	28.6M	21.0M

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure A.2. Post-trade returns relative to counterfactual, split by transaction size

This figure presents differences between average factor-neutral returns of stocks bought/sold and those of random buy/sell counterfactual strategies for buy and sell trades, grouped by transaction size. The bracket at the top of each bar is the 95% confidence interval of the point estimate at each horizon. Confidence intervals in brackets are computed using double-clustered standard errors, calculated as described in Appendix A.4. Transaction sizes are computed as a % of a fund's AUM, and broken into quintiles with break points computed by fund, using the full sample for each fund.



managers are expected to be stressed (high turnover trades) or which will have less immediate impact on the portfolio weights (small vs. large transactions). We also look at how the heuristic intensity results manifest, presenting evidence that they are less associated with a manager's average skill, but instead time-variation in their propensity to rely on extreme returns as a heuristic.

Figure A.2 presents results from an alternative approach in which we seek to separate selling decisions which are likely attended to from those that are not. Specifically, we compare the performance of large (top quintile, within portfolio) trades to all other trades. It is likely that portfolio managers think hard and attend to the largest trades, and in turn, we would expect those trades to perform better relative to a counterfac-

Table A.10. Buying and selling skill for subsamples with different levels of turnover

This table shows the factor neutral buying and selling vs. counterfactuals, but splitting the sample between funds with higher and lower turnover and different periods with higher and lower turnover. There is little difference between buying and selling skill for high and low turnover funds in the cross section, but a substantial deterioriation in selling skill relative to the counterfactual in periods of high turnover relative to a fund's typical (median) level.

	Panel A: Buys					**Panel B: Sells**				
	28	90	180	270	365	28	90	180	270	365
Low Turnover Funds	0.34	0.44	0.58	0.6	0.61	0.06	-0.06	-0.15	-0.33	-0.66
(below median)	(0.06)	(0.13)	(0.21)	(0.28)	(0.37)	(0.05)	(0.10)	(0.16)	(0.22)	(0.27)
High Turnover Funds	0.32	0.63	0.72	0.91	1.08	-0.09	-0.56	-1.25	-1.37	-1.21
(above median)	(0.07)	(0.12)	(0.18)	(0.25)	(0.28)	(0.07)	(0.18)	(0.29)	(0.37)	(0.43)
Low Turnover Periods	0.31	0.47	0.89	1.14	1.25	0.00	-0.04	-0.19	-0.18	-0.13
(below fund median)	(0.06)	(0.12)	(0.19)	(0.25)	(0.33)	(0.06)	(0.13)	(0.21)	(0.30)	(0.32)
High Turnover Periods	0.44	0.59	0.80	0.84	1.18	0.05	-0.28	-0.66	-0.77	-0.90
(above fund median)	(0.07)	(0.15)	(0.24)	(0.32)	(0.40)	(0.07)	(0.14)	(0.19)	(0.25)	(0.28)

tual. By contrast, if we believed the results to primarily be driven by a mechanical price impact, we would expect those same trades to perform the worst. Consistent with the attentional mechanism, all trade size quintiles other than the largest fail to beat the counterfactual. The median-sized trades underperform the counterfactual substantially.

Periods of high turnover are likely to be those during which a portfolio manager's attention is stretched thin. Prior work proposes that higher turnover is indicative of better trading opportunities, and thus higher future benchmark adjusted return, finding supporting evidence in mutual fund data (Pástor, Stambaugh, and Taylor 2017). While high turnover may indeed reflect better buying opportunities (buying performance measures are similar with higher gross buying volume), the mechanism outlined in the main text suggests that an asymmetric allocation of attention towards buying decisions would exacerbate the PM's tendency to sell poorly. Table A.10 shows that, consistent with such a mechanism, these periods are associated with particularly poor performance of selling decisions.

In the main text, we present evidence that the variation in heuristic intensity, using measures which sort across time within-manager, is linked with negative selling performance. We find that there is fairly limited variation in between manager differences in heuristics intensity. As one way of seeing this, we find that the overall result that heuristic intensity is related to selling underperformance is not sensitive to whether we compute within-fund sorts or across fund sorts. Table A.11 shows that we get similar results to the main text if we sort PMs into bins based on propensity to sell extremes at each point in time, suggesting that our results are not driven by persistent cross-sectional differences across funds in tendency to use heuristics but instead by time-variation in our heuristics proxy within-manager over time. This provides further evidence consistent with selling underperformance being related to an attentional mechanism rather than a persistent trait. In addition, we find that heuristic intensity is also largely uncorrelated with fund-level characteristics (for example, style or portfolio construction characteristics). These results are presented in A.12, which shows that there is no strong relationship between a fund's average style characteristics, concentration, turnover, or longer term focus and its average level of heuristic intensity.

A.9 Results using alternative currency adjustments

All of our analysis uses prices in USD, which implies a specific perspective on currency hedging. Because our sample has strategies which invest globally and have end clients in several countries other than the United States, this may not account for certain currency risks that managers are accounting for in their analysis.

To test this proposition, we calculate the modal country each strategy is invested in, and convert all of the factor-neutral security returns to that currency. Finally, we calculate the period returns for this portfolio, and average the Hold - Sell and Buy - Hold portfolios across time and strategies in the same manner as the USD based results in the paper. Results of this exercise are presented in table A.13.

A.10 Results from a calendar-time trading strategy

As discussed above, our baseline approach for assessing performance involves computing average returns across portfolios which are adjusted for systematic factors. In our baseline analysis, we double-cluster standard errors using the method described in Appendix A.4, which accounts for autocorrelation in contemporaneous trades of the same fund and across funds. Above, we find that the average stock sold outperforms the average stock held by about 80 bp over a one year horizon with a standard error of about 20 bp. The astute reader will note that such a performance differential would likely be very difficult to detect statistically if PMs were trading on many common

Table A.11. Post-trade sell returns relative to counterfactual by heuristics intensity, robustness checks

This table presents average returns relative to random sell counterfactuals for sell portfolios sorted by heuristics intensity, where the heuristics intensity measure is sorted across funds by week (panel I) and within funds over time (panel II). We divide these measures into four bins from Lowest, Low-Med, Med-High and Highest, based on their rankings. Columns represent factor-neutral sell performance measures at the following horizons: 1 month, 3 months, 6 months, 9 months, and 1 year. Heteroskedasticity and autocorrelation robust standard errors, computed using the method described in the section A.4, are reported in parentheses.

Heuristics	Bins	**Sell Results, Factor-neutral**				
Intensity (weekly)	Horizon	28	90	180	270	365
I. Across Funds,	Lowest	-0.07	-0.21	-0.29	-0.42	-0.28
by week		(0.07)	(0.14)	(0.20)	(0.24)	(0.33)
	Low-Med	0.00	-0.07	-0.16	-0.12	0.03
		(0.06)	(0.11)	(0.15)	(0.21)	(0.24)
	Med-High	0.05	0.05	-0.13	-0.20	-0.31
		(0.06)	(0.14)	(0.21)	(0.22)	(0.26)
	Highest	0.11	-0.59	-1.33	-1.57	-2.20
		(0.09)	(0.19)	(0.31)	(0.39)	(0.47)
II. Within Funds,	Lowest	-0.10	-0.30	-0.26	-0.30	-0.11
by week		(0.08)	(0.13)	(0.18)	(0.24)	(0.31)
	Low-Med	-0.04	-0.05	-0.27	-0.39	-0.24
		(0.06)	(0.10)	(0.13)	(0.23)	(0.26)
	Med-High	0.16	-0.02	-0.30	-0.37	-0.62
		(0.07)	(0.15)	(0.22)	(0.26)	(0.29)
	Highest	0.07	-0.50	-1.15	-1.34	-1.90
		(0.09)	(0.18)	(0.29)	(0.41)	(0.47)

systematic factors. As a result, our standard error calculations imply that one should be able to construct a calendar time strategy associated with buying stocks that PMs recently sold and taking short positions in stocks that were recently held which should have a high Sharpe ratio due to strong benefits from diversification. In this section, we will show that this is indeed the case, which provides an additional check on the robustness of our main results and inference.

Before describing our approach, we note our dataset is not ideally suited for this form of analysis. We have an unbalanced panel in which the number of portfolios changes over time, so such an approach implicitly overweights the portfolios which appeared early on in our sample. Second, since PMs in our sample operate in many

Table A.12. Cross-sectional correlations between fund average heuristic intensity measure and portfolio characteristics

This table shows the correlation between average heuristic intensity and average portfolio characteristics. The heuristic intensity (propensity to sell in extreme bins) is averaged over the full fund's history, as are the portfolio characteristics. This shows that correlations are near-zero, indicating that the average use of extreme returns by a fund is largely uncorrelated with a strategy's concentration, style, holding horizon, and turnover.

Characteristic	Cross-sectional Correlation
Concentration	0.02
Percent of Holdings held in Large Blocks	-0.01
Percent Long-Term Holdings	-0.02
Turnover	-0.02
Portfolio Average Book to Market Quintile	-0.05
Portfolio Average Size Quintile	-0.03
Portfolio Average Momentum Quintile	0.02

Table A.13. Post-trade returns relative to counterfactual, factor-neutral, hedged to modal portfolio currency

This table presents the average beta-neutral counterfactual returns for buy and sell trades hedged to each portfolio's modal currency. Standard errors are in parenthesis below the point estimates, calculated by the method described in Appendix A.4.

Horizon	28	90	180	270	365
Buy - Hold	0.24	0.40	0.59	0.57	0.63
	(0.05)	(0.09)	(0.13)	(0.18)	(0.21)
Hold - Sell	-0.11	-0.42	-0.74	-1.01	-1.05
	(0.05)	(0.11)	(0.17)	(0.22)	(0.26)

markets, there could be exposures to risk factors/style tilts that our current analysis differences out which are less easy to address by specifying constant loadings on a common set of risk factors. Ultimately, we believe that the extremely low degree of correlation across funds who trade at the same time (or even between stocks bought/sold by the same PM at the same time) should mitigate precisely these types of concerns that calendar time adjustments would normally be used to assuage. In contrast, we would expect a lot of correlation if they were trading on a set of common aggregate factors, though our estimates will find minimal exposures, in line with the similarity between our baseline factor-neutral and unadjusted performance results in Table 3.

In line with the standard approach of the literature, we conduct our analysis at the monthly frequency. To generate the portfolio, on a monthly basis, we construct signals based on the volume of securities sold and value of securities held on each given day by each given fund. The signal produced by each fund-day is the percent of total dollar volume made up by each security within a trading category. For example, if I sold \$5 worth of Apple, and overall I sold \$10 worth of securities, the fund/day weight on Apple within the sell signal would be 50%. In other words, for security i at time t in portfolio j and in transaction type $k \in buy, hold, sell$, we construct a signal as:

$$S_{k,i,j,t} = \frac{DollarVolume_{k,i,j,t}}{\sum_{i=1}^{N} DollarVolume_{k,i,j,t}}$$

where $DollarVolume$ is the price times the number of shares traded (or held, in which case we use portfolio weights).

This measure computed within each fund/day for each trade type. Days without sells are dropped from the counterfactual hold signal. The monthly signal for month m becomes

$$S_{k,i,m} = \left(\frac{1}{N_{port}}\right)\left(\frac{1}{N_m}\right)\sum_{j=1}^{N_{port}}\sum_{t=1}^{N_m} S_{k,i,j,t}$$

The short position of each strategy is given by the appropriate "hold" signal, and the long position of each is given by the "sell" signal, with leverage normalized to the equivalent of a 100% long-only weight on each side. In order to make sure that weights have equal leverage on each side, we normalize them in the following way:

$$S_{i,m}^{sell} = I(S_{sell,i,m} > S_{hold,i,m})(S_{sell,i,m} - S_{hold,i,m}),$$

$$S^{hold}_{i,m} = I(S_{hold,i,m} > S_{sell,i,m})(S_{hold,i,m} - S_{sell,i,m})$$

Since our signal predicts long-run underperformance, we construct 12 portfolios associated with the signal month, each at a different lag. We repeat this analysis, constructing portfolios which are associated with stocks sold minus stocks held, for various look-back periods. Specifically, we look at stocks which were traded/held between 1 and 12 months ago to form the portfolios. For example, given the signal in December 2009, we construct a January 2010 portfolio based on that signal, a portfolio which holds the same stocks with the same weights in February 2010, and so on through December.

To get an estimate most directly comparable to the one year cumulative returns discussed in the text, we average returns from the 12 portfolios formed with trading information from the prior 1 through 12 months. Figure A.3 shows the time-series of returns to this long-short strategy. The color of the line also indicates the number of portfolios included in the construction of trading signals at each point in time, which also shows one reason why we do not view this as the most appropriate way to analyze the data. As noted above, an important limitation in this analysis is that we have a somewhat unbalanced panel of strategy observations. We then average the twelve Sell - Hold portfolios, yielding a diversified portfolio associated with 1 year worth of past signals.

As is clear from the graph, the strategy does quite well. It has both positive returns and low volatility, and improves as we see a larger number of portfolios enter our sample. Reassuringly, the performance of the strategy is best during periods of time when our signals are aggregated over the largest number of funds.

As is clear from the steep slope and low volatility of the graph, this portfolio has a very large Sharpe ratio, with both low volatility and high returns (though the numerator is largely a function of how we normalize the weights). Table A.14 computes the Sharpe ratio associated with our one year portfolio. In order to make sure that different systematic factors are not driving the results, we adjust for a variety of factor models: the CAPM, the Fama-French 3 factor model, and the Carhart 4 factor model and report Sharpe ratios associated with portfolios which are hedged with respect to these factors.

We also perform the same analysis with different formation windows, to show that the specific choice of one year is not driving the results. The portfolios based on the first 6 months after trade have almost the same Sharpe ratios as those based on the

Figure A.3. Cumulative returns for a trading strategy which purchases stocks recently sold and sells short stocks which were recently held over the prior year

This figure reports the time series of cumulative monthly returns associated with a trading strategy which takes long positions in stocks which were disproportionately sold over the prior year and short positions in stocks which were recently held over the same periods. Note that our sign convention differs from the rest of the paper, so positive numbers indicate that sells outperform stocks held (which acts as a drag on performance). Color of the line indexes the number of strategies we observe in a given month over the course of our full sample period. Please see the text for further details on portfolio construction.

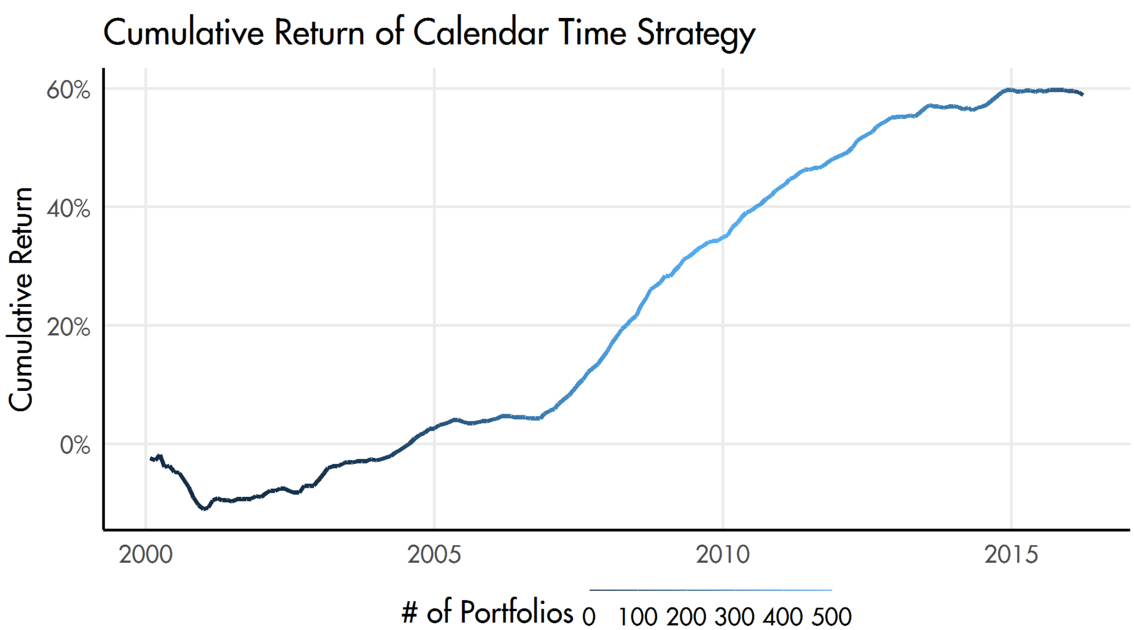


Table A.14. Sharpe Ratios for a variety calendar time portfolios which purchase stocks recently sold and sell short stocks which were recently held

This table reports annualized Sharpe ratios associated with a number of long-short strategies based on a signal constructed from the previous sells and holds of the managers in our sample. In the first row, we report raw Sharpe ratios. In subsequent rows, we report Sharpe ratios that obtain after hedging exposures to different factors. In each case, we average over multiple portfolios which are formed based on trading activity from one or more month ago. Columns indicate which periods are averaged when computing the overall portfolio. Signals are constructed based on the percent of daily volume a security made up in the sell portfolio, and based on the portfolio weight relative to other holds for the hold portfolio (please see text for further details on portfolio construction). Sharpe ratios are computed assuming that stocks recently sold are purchased for the long end and the stocks recently held are sold short.

	1 Year	Months 2:12	Months 1:6	Months 7:12
No Adjustment	1.496	1.312	1.037	1.150
CAPM	1.484	1.296	1.030	1.138
Three Factor	1.669	1.469	1.141	1.260
Four Factor	1.671	1.467	1.130	1.276

last 6 months. In addition, we exclude the first month due to concerns about trading costs, and demonstrate that the portfolio does about as well without include the first month following each "sell" observation. The relationship between the selling signal and future performance is fairly persistent, so the lower denominator tends to dominate in the Sharpe ratio calculations, leading them to increase as we aggregate over longer periods.

Table A.15 reports alphas and factor loadings for these portfolios, the latter of which are almost always close to zero, indicating that this signal is largely unrelated to potential known risk factors which are related to returns. These results underscore our findings from above suggesting that PMs in our sample appear to be trading on individual stocks rather than seeking systematic exposures to common factors.

In closing, we note that actually implementing this trading strategy would likely result in lower returns than we see here because of the high transaction associated with a high turnover strategy investing in large numbers of securities globally. It is also not likely to be tradeable in practice because of the need for daily holdings information across a variety of skilled managers. As such, we view this exercise as largely supporting

Table A.15. Abnormal returns and factor loadings for trading strategy which purchases stocks recently sold and sells short stocks which were recently held over the prior year

This table reports annualized Sharpe ratios, along with monthly alphas and factor loadings associated with a trading strategy which takes long positions in stocks which were disproportionately sold over the prior year and short positions in stocks which were recently held over the same periods. We report measures for monthly performance for a number of factor specifications. Robust standard errors are in parentheses. Please see the text for further details on portfolio construction.

	Ann. Sharpe	Monthly Alpha (%)	Mkt - RF	SMB	HML	Mom
No Adjustment	1.496	0.238				
		(0.040)				
CAPM	1.484	0.236	0.00702			
		(0.0418)	(0.011)			
Three Factor	1.669	0.257	0.0128	-0.0377	-0.0342	
		(0.0383)	(0.0113)	(0.0167)	(0.0187)	
Four Factor	1.671	0.257	0.0124	-0.0375	-0.0344	-0.000769
		(0.0389)	(0.0127)	(0.0165)	(0.0181)	(0.0104)

our baseline result and providing another perspective on the robustness of our standard error calculations.