Journal of Financial Economics 137 (2020) 515–549

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

What you see is not what you get: The costs of trading market anomalies☆

Andrew J. Patton*, Brian M. Weller

Duke University, 213 Social Sciences Building, Box 90097, Durham, NC 27708, USA

ARTICLE INFO

Article history:
Received 28 November 2018
Revised 19 April 2019
Accepted 1 May 2019
Available online 5 March 2020

JEL classification:
G12
G14
G23

Keywords:
Trading costs
Performance evaluation
Mutual funds
Market efficiency

ABSTRACT

Is there a gap between the profitability of a trading strategy on paper and that which is achieved in practice? We answer this question by developing a general technique to measure the real-world implementation costs of financial market anomalies. Our method extends Fama-MacBeth regressions to compare the on-paper returns to factor exposures with those achieved by mutual funds. Unlike existing approaches, ours delivers estimates of all-in implementation costs without relying on parametric microstructure models or explicitly specified factor trading strategies. After accounting for implementation costs, typical mutual funds earn low returns to value and no returns to momentum.

1. Introduction

Empirical asset pricing overflows with explanations for differences in average returns across securities. The proliferation of predictors distracts from genuine market anomalies from which lessons could be drawn about risks, preferences, and beliefs. Recent calls to action by Harvey et al. (2016), Harvey (2017), and Hou et al. (2017) have focused on high false discovery rates and scurrilous academic practices. Fundamentally, they question whether candidate factors in the cross section of expected returns are real and actionable.

We give on-paper trading strategies the benefit of the doubt and instead investigate whether they are implementable in practice, thereby representing true expected return factors or market anomalies. This line of inquiry originates with Fama (1970), who considers the role of transactions costs in defining market efficiency and departures therefrom.

Despite nearly 50 years of subsequent research, accurately measuring real-world implementation costs for academic factors remains a formidable challenge. Existing approaches generally fall into two categories. The first category entails using proprietary trading data to analyze costs for a single firm (e.g., Keim and Madhavan, 1997; Engle et al., 2012; Frazzini et al., 2015). Although selected firms may not be representative of asset managers as a whole, such analyses provide an informative lower bound on the

☆ We thank Bill Schwert (Editor), Eugene Fama (referee), Mark Carhart, Alex Chinco, Andrew Detzel, Matthew Farrell, Andrea Frazzini, Campbell Harvey, David Hsieh, Robert Korajczyk, Alex Kostakis, Juhani Linnainmaa, Toby Moskowitz, Ronnie Sadka, Michela Verardo, and seminar participants at Duke University, University of Manchester, University of Warwick, the American Finance Associate annual meeting, the Imperial College Annual Hedge Fund Conference, the Rodney L. White Center Conference on Financial Decisions and Asset Markets, the Yale School of Management Junior Finance Conference, the University of British Columbia Winter Finance Conference, and John Cochrane's 60th Birthday Conference for valuable comments and discussion.

* Corresponding author.
E-mail addresses: andrew.patton@duke.edu (A.J. Patton), brian.weller@duke.edu (B.M. Weller).

https://doi.org/10.1016/j.jfineco.2020.02.012
0304-405X/© 2020 Published by Elsevier B.V.

implementation costs of factor strategies. The second approach uses market-wide trading data such as NYSE Trade and Quote (TAQ) to estimate trading costs for individual securities and then cumulate simulated costs of trade implied by dynamic factor strategies (e.g., Lesmond et al., 2004; Korajczyk and Sadka, 2004; Novy-Marx and Velikov, 2016). Papers in this category typically extrapolate price impact estimates from small trades to large factor portfolios or ignore price impact costs entirely.

Our work complements these approaches with a new cross-sectional technique that combines the best elements of both. Like papers that utilize proprietary trading data, our estimates reflect the all-in costs of implementing factor strategies, and they apply equally well for past and modern market environments [for which the zero-return day measure of Lesmond et al. (1999) fails, for example]. Like papers that estimate transaction cost functions using market data, our baseline methodology captures the costs faced by *representative* practitioners of factor investing instead of single, special investment managers. In contrast with both approaches, our methodology facilitates the evaluation of implementation costs without specifying the precise trades used to implement factor strategies and for arbitrary subsets of the asset management universe. These innovations are important because existing methods using precisely specified factor strategies with different data sources and sets of firms disagree on the implementability of factor strategies. For example, Lesmond et al. (2004) find no net-of-costs return to momentum using TAQ data and a representative set of traders, whereas Frazzini et al. (2015) find positive momentum premia for a large hedge fund. Although our methodology rests on few assumptions, it has sufficient resolution to reconcile these disparate results in a transparent way.

Our methodology is an extension of the familiar Fama and MacBeth (1973) procedure. In the first stage, time series regressions estimate factor loadings β_i for each asset *i*, and in the second stage, cross-sectional regressions estimate the compensation per unit of factor exposure λ_t at each date *t*. Standard assets are based on stock portfolios, and the resulting estimates of compensation for factor exposure, denoted $\bar{\lambda}_k^S$, represent the on-paper profitability of a given factor strategy. We augment the set of assets to include 4267 US domestic mutual funds, and we allow the compensation for factor exposure earned by mutual funds, $\bar{\lambda}_k^{MF}$, to differ from that which is available on paper. Unlike stock portfolio returns, (gross) mutual fund returns reflect the real-world implementation costs of factor strategies.[1] Thus, the difference between mutual fund and stock portfolio compensation $\bar{\lambda}_k^S - \bar{\lambda}_k^{MF}$ delivers an estimate of implementation costs for factor *k*.[2] Because costs per unit of exposure are likely to be negatively correlated with factor exposures, that is, funds that earn greater net returns to a factor are more likely to take greater exposures to it, our estimate of implementation costs represents a lower bound on the costs faced by a representative mutual fund.

Our empirical analysis focuses on the implementation costs of mutual funds for the market (*MKT*), value (*HML*), size (*SMB*), and momentum (*UMD*) factors. We choose these factors because they comprise the dominant empirical models in academic finance (e.g., Fama and French, 1992; Carhart, 1997) and because they serve as the basis for hundreds of billions of dollars in investments. We study mutual funds as our set of asset managers because they collectively manage more than $16 trillion of capital in the United States, and the mutual fund industry has been better populated for a longer period of time than alternative asset managers such as hedge funds.[3] Our approach is readily extended to other factors and market participants.

Our analysis delivers two new empirical facts on the all-in implementation costs of anomalies for typical mutual funds. First, momentum strategies suffer extreme underperformance in practice relative to on-paper strategies. Our full-sample estimates of all-in implementation costs are in the range of 7.2%–7.6% per year, which eliminates most profits accruing to momentum during the 1970–2016 period. About half of this cost is due to mutual funds' inability to short. Our all-in cost estimates are considerably larger than those typically estimated using bid-ask spreads alone (e.g., Novy-Marx and Velikov, 2016). We conclude, as Lesmond et al. (2004) do, that momentum strategies are unprofitable for typical asset managers when a broader set of implementation costs are considered. Second, mutual fund implementation costs sharply reduce returns to the value factor; we estimate all-in costs of 2.6%–4.1% per year. In contrast, mutual funds implementation costs for the market and size factors are approximately zero.

Our approach also yields insights into the sources of implementation costs for typical firms. Simple modifications to the set of test portfolios, factors, and slopes considered allow us to attribute costs to three primary sources. First, by excluding microcap stock portfolios, we can gauge the potential shadow costs of investability restrictions faced by real-world investors. Doing so, we find that difficulty in investing in the smallest stocks explains reductions in realizable factor compensation of 1% per year for value and momentum. Second, although mutual funds face shorting constraints, it is common to use the familiar long/short portfolios small-minus-big (SMB), high-minus-low (HML), and up-minus-down (UMD) as factors in asset pricing regressions.[4] We develop two long-only variants of the Carhart factors to assess the role of institutional constraints on shorting. We find that shorting frictions explain roughly half of mutual fund under-performance on momentum and between one-fifth and one-third of under-performance on value. Third, we consider the gap attributable to mutual funds tracking alternative variants of the usual academic factors. Sorting funds by their time series R^2 from the four-factor model, we estimate that around half of the average mutual fund under-performance

[1] We use gross returns to focus on the efficiency of mutual funds' investing technology instead of on the distribution of rents between managers and investors embedded in net returns.

[2] Our more sophisticated approaches account for time- and cross-sectional variation in implementation costs, which we discuss further below.

[3] See the "2017 Investment Company Fact Book," available at http://www.icifactbook.org/.

[4] On shorting constraints, note that the Investment Company Act of 1940 explicitly caps leverage for most mutual funds at one-third.

on value and momentum is associated with uncompensated departures from the academic factors.[5]

As a third empirical contribution, we analyze variation in implementation costs across funds and time and demonstrate the importance of considering such variation in evaluating the implementability of factor strategies. While the typical firm's compensation for momentum is indistinguishable from zero, subsets of the mutual fund universe can achieve positive returns to momentum net of costs. A focused analysis on smaller subsets is also important from an aggregate market efficiency perspective because a violation exists if some investors see anomalous profits, even if a typical investor does not. For this purpose, we segment the mutual-fund universe by (lagged) total net assets. Size is a natural sorting dimension because Berk and Green (2004), Pastor et al. (2015), Berk and van Binsbergen (2015), and others link scale to gross-of-fees performance. We rerun our cross-sectional analysis using each mutual fund size category separately, and we confirm that small and large mutual funds achieve different returns to momentum from typical mutual funds. Using this insight, we reconcile conflicting evidence on the transactions cost rationale for the continued existence of the momentum anomaly.[6]

Our approach provides an estimate of the gap in factor-mimicking portfolio performance ($\lambda_{kt}^{S} - \lambda_{kt}^{MF}$) for each particular factor and date, and we use this information to study determinants of the time series of average implementation costs. We show that industry inflows are associated with increased strategy costs, which in turn neutralize the secular declines in bid-ask spreads that affect the first dollar traded in factor strategies. As a consequence, bid-ask–spread based measures increasingly underestimate the true costs of factor strategies as asset management (and factor investing in particular) grows in scale.

While our new approach delivers simple, nonparametric, estimates of the implementation costs for factor trading strategies, it does face some limitations. First, as mentioned above, our approach delivers lower bounds on implementation costs. In our empirical analysis these bounds do not greatly limit the economic conclusions we can draw: the estimated costs are already so high as to eliminate or severely attenuate the on-paper profitability of strategies such as value and momentum for typical mutual funds. For other strategies, estimates that indicate positive returns net of costs do not necessarily imply that an anomaly can be implemented by typical investors. In this sense, our measures can diagnose an implementability problem with a factor, but they cannot deliver a clean bill of health.

Secondly, our technique relies on real-world asset managers to reveal implementation costs through realized returns to their chosen factor exposures. We cannot speak to the costs of new factors that asset managers have not had an opportunity to trade.[7] For the same reason, our approach cannot estimate implementation costs for counterfactual factor exposures to evaluate strategy carrying capacities, unlike approaches that rely on parametric transaction cost models.

Finally, like much of the literature on performance evaluation, our method is susceptible to criticism of the choice of factors included in the analysis. A manager who is following a strategy that does not correspond to an approximate linear combination of those included in the model could appear to have high implementation costs for the included strategies, even though she has low costs for the strategy actually being implemented. We verify that omitted mutual fund strategies do not drive our high implementation cost estimates by replicating large performance gaps for funds with returns almost completely explained by the academic factors (the average R^2 of the four-factor model for these funds' return histories is 94%). For these funds, the scope for omitted strategies is too small to explain the observed real-world performance gaps.

The remainder of the article is structured as follows. Section 2 relates our work to existing papers in the literature. Section 3 describes our data. Section 4 presents our findings on the implementation costs faced by mutual funds. Section 5 decomposes estimated implementation costs into components due to shorting constraints, investment restrictions, and tracking errors. Section 6 examines implementation costs across different types of funds and across time. Section 7 concludes. The appendix contains additional details and analyses.

2. Related literature

The Fama and French three-factor model has been the benchmark for empirical asset pricing since its introduction in 1992. This empirical model supplanted the capital asset pricing model (CAPM), but its new value and size factors had little theoretical motivation.[8] As factors continued to emerge over the next quarter century, most notably, the momentum anomaly of Jegadeesh and Titman (1993), several strands of literature emerged in an attempt to tame the "factor zoo" (Cochrane, 2011). One active strand investigates the implementation costs of anomalies with a particular focus on size, value, and momentum anomalies. While implementation costs cannot explain why expected return discrepancies come to be in the first place, this literature seeks to rationalize the continued existence of market anomalies as their by-product. Our paper advances this line of inquiry by introducing a new and readily generalizable approach for measuring the real-world implementation costs of return factors and anomalies.

[5] This analysis also addresses a potential concern about the strategies that mutual funds trade, which is discussed below.

[6] We also run subsample analyses by quintile of total net assets and four-factor R^2s. Our methodology can accommodate many other splits of interest, e.g., sorting by factor betas sheds light on typical gains to running combined strategies. We leave investigations of other cuts of the mutual fund universe to future work.

[7] This caveat does not apply in the particular case of momentum. Grinblatt et al. (1995) argue that momentum-like strategies are endemic among mutual funds in their 1975–1984 sample, decades before the publication of Jegadeesh and Titman (1993).

[8] Banz (1981) and Basu (1977) show price-earnings ratios and market capitalization as characteristics associated with deviations from the CAPM.

Existing methods for measuring implementation costs take two approaches. The first approach uses specialized trading data to evaluate the costs of trade for large investment managers with the implicit assumption that these managers are representative of sophisticated investment managers more generally. These papers typically assess trading costs using the Perold (1988) implementation shortfall measure, which captures the difference between realized profits and on-paper profits using a preset decision price. This approach dates back at least to the Keim and Madhavan (1997) analysis of the transactions costs of a variety of investment styles for $83 billion of trades.[9] In this vein, Keim (2003) uses institutional trading data for 33 firms and finds that trading costs likely eliminate profits to on-paper momentum strategies.

A key challenge to this method is that institutional trading is endogenous: traders are particularly aggressive in their trading targets when liquidity is readily available, which in turn imparts a downward bias to estimated cost functions. Frazzini et al. (2015) overcome this challenge by using data from an investment manager whose trading targets are model-generated and selected irrespective of market conditions. Armed with more than $1 trillion of trades, they analyze value, size, and momentum anomalies and find that all of them are implementable and scalable to tens or hundreds of billions of dollars of invested capital. By their reckoning, and in contrast to the Keim (2003) managers, major anomalies continue to be anomalous if their asset manager's costs are representative of typical investment managers' costs.

The second approach trades off accuracy for representativeness in estimating implementation costs. Instead of using proprietary trading data for a single asset manager to estimate costs directly, other studies derive transactions costs using aggregate price and transaction records and extrapolate estimated price impact functions to factor trading strategies.[10] Much of this literature focuses on the momentum anomaly because of its high turnover, and even the originating article establishing the momentum anomaly considers a trading costs explanation (see Jegadeesh and Titman, 1993, and later Jegadeesh and Titman, 2001). Notably, none of these papers use precise all-in trading cost measures like implementation shortfall because theoretical or decision-date prices are not obtainable outside of specialized trading data.

Chen et al. (2002) estimate separate price impact functions for 5173 individual stocks and calculate the trading costs accruing to size, value, and momentum strategies. The authors suggest that all factors have break-even carrying capacities on the order of millions of dollars (*HML*) to hundreds of millions of dollars (*SMB*). By their calculations, factor strategies are not investable. Lesmond et al. (2004) suggest that momentum in practice trades in "disproportionately high cost securities" rather than the typical transactions cost securities Jegadeesh and Titman (1993) use for approximating the momentum factor's trading costs. Using effective spreads from TAQ, commission schedules from a discount brokerage, and all-in frictions implied by zero-trading days (Lesmond et al., 1999), Lesmond et al. (2004) argue that trading costs erase the returns to the momentum anomaly.

Korajczyk and Sadka (2004) present more optimistic results on the investability of factor strategies. Korajczyk and Sadka (2004) use TAQ data to estimate effective and quoted spreads, the primary proportional costs studied in the literature, and price impact or "non- proportional trading cost" functions from Glosten and Harris (1988) and Breen et al. (2002). In utilizing different non-proportional cost functions from Lesmond et al. (2004), Korajczyk and Sadka (2004) extrapolate trade-level costs to find positive net-of-cost returns to the momentum anomaly. They invert their cost function estimates to obtain a break-even momentum strategy carrying capacity of $5 billion. Novy-Marx and Velikov (2016) measure trading costs using effective spreads recovered from the Hasbrouck (2009) Bayesian Gibbs sampler and tally costs of trading size, value, and momentum strategies, among others. The authors estimate strategy carrying capacities of $5 billion for momentum (as in Korajczyk and Sadka, 2004), $170 billion for size, and $50 billion for value, the latter two of which are comparable to the estimates of Frazzini et al. (2015). These approaches do not account for the price impact costs of large institutional investors, and they likely overestimate the true strategy carrying capacities as a result.

In concurrent work, Arnott et al. (2017) argue, as we do, that mutual funds deliver much lower returns on value and momentum anomalies than on-paper factor counterparts could indicate. Our paper differs from theirs in four key respects. First, we modify standard Fama-MacBeth regressions to develop a cost estimation procedure that is robust to heterogeneity in implementation costs across both funds and time. Second, we decompose costs to highlight the respective roles of shorting, investability, and liquidity frictions. Third, we slice the cross section of mutual funds to distinguish among funds of different attributes and in so doing reconcile previous work on implementation costs. Fourth, our approach compares factor-mimicking portfolio returns for mutual funds and stock portfolios. The Arnott et al. (2017) use of on-paper factor returns as a benchmark is valid only if investors can frictionlessly trade stock factors.

3. Data

Our mutual fund sample consists of 4267 United States domestic equity mutual fund groups with at least 24 nonmissing monthly gross returns from January 1970 to December 2016. In Appendix A, detailing our mutual fund filtering methodology, we describe a number of data cleaning and filtering steps based on the recommendations of Berk and van Binsbergen (2015), Pastor et al. (2015), and others. One data processing step bears special mention: we map funds delineated by share class into fund groups. Share classes for funds with identical investments differ

[9] Other studies use the Keim and Madhavan (1997) calibrated transaction cost functions to decompose fund performance for a larger universe of funds. For example, Wermers (2000), like our study, finds that implementation costs meaningfully erode mutual fund returns.

[10] Grundy and Martin (2001) and Barroso and Santa-Clara (2015) invert this logic and calculate the transactions costs that would be required to wipe out the momentum anomaly.

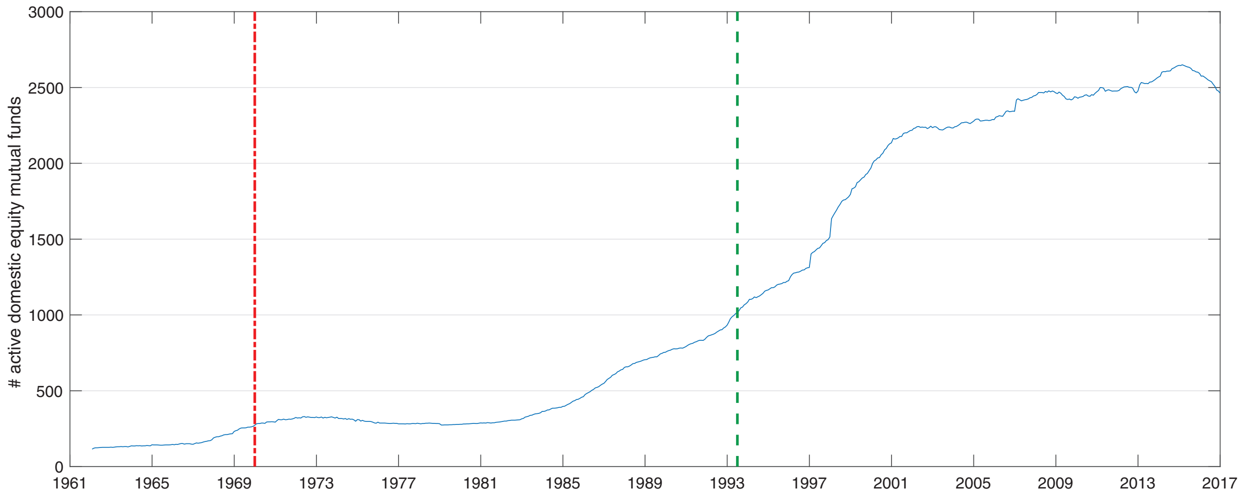

Fig. 1. Count of active domestic equity mutual funds by month.
This figure plots the count of non-missing returns by month for United States domestic equity mutual funds. The dashed line at January 1970 marks the starting point of our 1970–2016 sample. The dashed line at July 1993 marks the midpoint of the post-1970 sample as well as the start date for our post–Jegadeesh and Titman (1993) sample.

in fees charged to investors, but they are not otherwise economically distinct. To aggregate returns within a fund group, we take total net asset (TNA) weighted gross-of-fee returns. The Center for Research in Security Prices (CRSP) provides returns net of management and 12b-1 fees, and we convert these into gross returns by adding expense ratios divided by 12, following Fama and French (2010). We use "fund group" and "fund" interchangeably henceforth.

Significant changes in the count of active mutual funds reflect both a secular growth in the mutual fund industry and continual improvements in data quality.[11] Fig. 1 highlights these changes by plotting the number of non-missing returns for domestic equity mutual funds by month. The number of funds increases from 276 in January 1970 to 979 in July 1993 to 2463 in December 2016. Because the number and composition of funds varies widely over time, we conduct our analysis both on an extended sample and on a more recent subsample. Our long sample runs from January 1970 to December 2016. We discard the 1962–1969 window during which monthly returns are less consistently provided and during which several of our liquidity proxies are not available. Our recent subsample consists of the second half of the long sample and runs from July 1993 to December 2016. This start date postdates the Jegadeesh and Titman (1993) evidence of the momentum anomaly, the most recently discovered factor we consider. Table 1 reports summary statistics for the set of mutual funds used in our analysis. All told, the 1970–2016 sample consists of 724,995 fund-month observations and the 1993–2016 sample consists of 597,992 fund-month observations.

Much of our analysis compares mutual funds with similar stocks as measured by loadings on equity risk factors. Our Fama-MacBeth tests of Section 4 combine mutual fund data with common test portfolios. Our first portfolio set consists of the Fama-French 25 size-value double-sorted portfolios plus 25 size-beta portfolios, 25 size-prior return portfolios, and 25 size-Amihud illiquidity portfolios to ensure adequate dispersion in factor loadings to identify risk premia. We supplement this set of test assets with an expanded cross section following the recommendation of Lewellen et al. (2010). In our larger portfolio set, we add 49 industry portfolios, 25 size–operating profitability portfolios, 25 size–investment portfolios, 10 market beta–sorted portfolios, 10 market capitalization–sorted portfolios, 10 book-to-market ratio–sorted portfolios, 10 prior-return–sorted portfolios, 10 Amihud illiquidity–sorted portfolios, 10 operating profitability–sorted portfolios, and 10 investment–sorted portfolios for a total of 269 portfolios. With the exception of the illiquidity-sorted portfolios, all portfolio data are downloaded from Ken French's website.[12] Decile illiquidity portfolios sort stocks by the median daily Amihud illiquidity (daily absolute returns over dollar volume) over the prior calendar year, and stocks are assigned for the following year using deciles computed at the end of June to match the timing convention of the other portfolio data.[13] The 25 size–illiquidity portfolios sort first on lagged market capitalization and then on Amihud illiquidity quintile within each size bin to ensure that all portfolios are nonempty. Our analysis uses both equal- and value-weighted stock portfolios.

[11] Pages 1–2 of the CRSP mutual fund database guide details the amalgamation of data sources used to construct returns from December 1961 through the present. Page 16 discusses the merge of classifications into CRSP objective or style codes that we use to restrict the set of funds to United States domestic equity funds.

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

[13] Our monthly stock sample consists of all CRSP stocks (share codes 10 or 11) with at least 24 non-missing monthly returns, for a total of 22,121 unique PERMNOs over the 1970–2016 sample period.

Table 1
Domestic equity mutual fund sample summary statistics.
This table presents summary statistics for the 1970–2016 sample of 4267 United States domestic equity mutual funds. Panel A provides information on the time series of the number of active funds for each date as well as cross-sectional information on fund lifetimes and total net assets (TNA) at sample start, middle, and end. Panel B reports distributional information on fund excess returns, such as the mean return, return volatility and the Sharpe ratio. $\bar{\rho}_{MF}$ is the average pairwise correlation with other mutual funds' returns, and $\rho_{S\&P500}$ is the correlation with the S&P 500 index return.

Panel A: Mutual fund counts, ages and sizes

	Funds (number)	Lifetime (years)	TNA, Jan. 1970 (Million USD)	TNA, July 1993 (Million USD)	TNA, Dec. 2016 (Million USD)
Mean	1286	14.16	128.74	552.87	2590.70
Std. Dev.	917	10.50	302.83	1533.70	13254.00
25%	324	5.75	3.96	37.48	70.93
50%	1023	11.58	23.90	118.36	314.00
75%	2282	19.58	91.18	431.83	1421.30

Panel B: Mutual fund return distribution characteristics

	Mean Return (%/Month)	Return Vol. (%/Month)	Sharpe ratio (Annualized)	$\bar{\rho}_{MF}$ (%)	$\rho_{S\&P500}$ (%)
Mean	0.46	4.88	0.41	74.10	84.97
Std. Dev.	0.63	1.97	0.41	16.49	18.42
25%	0.32	3.86	0.25	71.77	81.94
50%	0.56	4.66	0.44	77.86	89.55
75%	0.78	5.59	0.60	82.32	94.53

We include several market and funding liquidity variables to proxy for time-varying cost factors that can affect the performance of mutual funds relative to stocks. Our market liquidity variables are Amihud illiquidity (Amihud, 2002), Pastor-Stambaugh liquidity levels (Pastor and Stambaugh, 2003), NYSE-average bid-ask spreads (Corwin and Schultz, 2012), and the Chicago Board of Exchange Standard & Poors 500 (CBOE, S&P 500) Volatility Index (VIX), as motivated by Nagel (2012). We use the Corwin and Schultz (2012) methodology to estimate bid-ask spreads because it enables measurement of market liquidity before TAQ becomes available in 1993 and because it captures average effective spread levels and innovations better than other pre-TAQ methodologies (see Corwin and Schultz, 2012, Table IV).[14] We use the CBOE S&P 100 Volatility Index (VXO) in place of the VIX in the pre-1990 period for which the VIX is not available. We compute Amihud illiquidity using CRSP daily data with values averaged within a month as in Amihud (2002), and we obtain the Pastor-Stambaugh series and CBOE VXO and VIX series from Robert Stambaugh's website[15] and the Federal Reserve of St. Louis's FRED database, respectively.

Our funding liquidity variables are the Frazzini and Pedersen (2014) "betting against beta" (BAB) factor, the He et al. (2017) intermediary capital ratio, the ten-year BAA minus ten-year Treasury spread, and the three-month London Interbank Offered Rate (LIBOR) minus three-month Treasury yield or "TED" spread. The first two series are expressly designed to capture institutions' funding liquidity constraints, and the latter two series are common proxies in the funding liquidity literature (e.g., Brunnermeier, 2009). We obtain BAB from AQR's website,[16] intermediary capital ratios from Asaf Manela's website,[17] and credit and TED spreads from FRED.

4. Fama-MacBeth estimates of implementation costs

4.1. Fama-MacBeth methodology

In this section, we consider the compensation per unit of risk exposure and investigate whether mutual funds obtain the same risk premium that academics achieve on paper. In our baseline estimation, we assume that mutual funds have a constant per-unit cost for implementing academic anomalies. Investing in a market index with $\beta_{MKT} = 1$ results in a performance gap of η relative to the on-paper performance of a market index, and investing in a levered version of the market more generally results in a performance gap of $\eta\beta_{MKT}$. In this setting, we would expect performance differences between stock and mutual fund portfolios to be linear in factor exposure.

We estimate the "implementation gap" using augmented Fama and MacBeth (1973) two-stage regressions for the Carhart four-factor model (Carhart, 1997). The time series regression step is standard except for the choice of test assets. We have $N_S = 100$ and $N_S = 269$ stock portfolios for the baseline and extended portfolio sets, respectively. In addition to stock portfolios, we $N_{MF} = 4267$ mutual funds, of which more than a thousand are active in the typical month. As diversified entities spanning a wide range of multifactor risk exposures, mutual funds,

[14] Corwin and Schultz make their code available at https://www3.nd.edu/~scorwin/HILOW_Estimator_Sample_002.sas. As in their paper, we compute cross-sectional averages using only NYSE-listed stocks, and we use their variant of estimated spreads in which negative values are set to zero.

[15] http://finance.wharton.upenn.edu/~stambaug/

[16] https://www.aqr.com/Insights/Datasets/Betting-Against-Beta-Equity-Factors-Monthly

[17] http://apps.olin.wustl.edu/faculty/manela/data.html

unlike stocks, need not be grouped into portfolios via a characteristic-sorting procedure.

The $N_S + N_{MF}$ first-stage time series regressions are

$$r_{it} = \alpha_i + \sum_k f_{kt}\beta_{ik} + \epsilon_{it}, \quad i = 1, \ldots, N_S, N_{S+1}, \ldots, N_S + N_{MF}, \tag{1}$$

where r_{it} is the month t gross return on stock portfolio or mutual fund i net of the contemporaneous risk-free rate and f_{kt} (for $k = 1, \ldots, K$) is the return on factor k at date t. The usual second-stage cross-sectional regressions are extended to accommodate the possibility of differences in risk pricing for stocks and mutual funds,

$$r_{it} = \sum_k \lambda^S_{kt}\hat{\beta}_{ik}1_{i\in S} + \sum_k \lambda^{MF}_{kt}\hat{\beta}_{ik}1_{i\in MF} + \epsilon_{it}, \quad t = 1, \ldots, T. \tag{2}$$

The regression in Eq. (2) is equivalent to two separate cross-sectional regressions run on stocks and mutual funds because the indicators partition the set of observations and coefficients. λ^S_{kt} and λ^{MF}_{kt} represent the factor-mimicking portfolio returns for stocks and mutual funds, that is, the hypothetical date t returns to a stock or mutual fund portfolio with $\beta_k = 1$ and $\beta_j = 0\,\forall\, j \neq k$.[18] If these factors were tradeable by real-world investors, f_{kt}, λ^S_{kt}, and λ^{MF}_{kt} would all be equal. The difference $\hat{\lambda}^{\Delta}_{kt} \equiv \hat{\lambda}^S_{kt} - \hat{\lambda}^{MF}_{kt}$ is our estimate of the implementation costs for strategy k, and it is the gap between the on-paper returns of a given strategy ("what you see") and the actual returns achieved by an asset manager facing real-world implementation costs ("what you get"). Conceptually this difference captures both direct costs such as spreads and price impact from factor trading as well as indirect costs such as investing in illiquid versions of factors to robustify strategies against outflows. Our point estimates are the average of the monthly differences in factor compensation $\bar{\lambda}^{\Delta}_k$, and we construct Newey and West (1987) standard errors for this difference using three monthly lags to account for serial correlation and heteroskedasticity in the λ-difference series.

Throughout our analysis, we estimate cross-sectional slopes of returns on risk exposures assuming that risk exposures are constant. In making this assumption, we prioritize minimizing the errors-in-variables problem arising from using noisy betas as inputs in the second-stage Fama-MacBeth regression. This problem is vitally important because we do not want to find differences in λs simply as a by-product of higher measurement error in mutual fund betas. Static betas effectively eliminate this issue. Appendix F confirms that measurement error has little effect on our results using univariate estimates of potential attenuation effects and a new, instrumental-variables approach to correct for measurement error in betas (Jegadeesh et al., 2019). In using static betas, we trade off against taking on model misspecification arising from time-varying stock portfolio or mutual fund risk exposures. Note, however, that if funds on average have timing ability, then using static betas in place of time-varying betas understates true implementation costs. For example, if funds scale up their betas when λ is high, then cross-sectional slopes for mutual funds $\hat{\lambda}_t$ are biased up, and the average estimated factor compensation $\bar{\hat{\lambda}}$ exceeds its true value.

Following Lettau et al. (2014) and others, we omit the constant term in Eq. (2) to force cross-sectional average alphas to zero. Economically, this omission forces the typical zero-risk security or mutual fund to have zero excess (gross) return at each point in time. We impose this restriction because the slope on β_{MKT} is not otherwise well identified in our stock portfolio sample; that is, the time series of the intercept α_t and the estimated market risk premium $\lambda_{MKT,t}$ are strongly negatively correlated and of similar magnitudes. By contrast in the mutual fund sample, market beta has a large and positive risk price regardless of whether a constant is included. Empirically, the dominant effect of suppressing the constant is delivering a reasonable estimated market risk premium for stock portfolios, and no factor risk premia are meaningfully affected for mutual funds.[19]

4.2. Baseline estimates

Table 2 presents estimates of Eq. (2). The λ^{Δ} value in the upper-left corner indicates that the difference in compensation per unit of market exposure is 0.38% per year greater for risk exposures taken via mutual funds than in (one hundred value-weighted) on-paper stock portfolios. This difference declines slightly to 0.21% per year when assessed against the full set of 269 portfolios. Neither effect is statistically or economically significant, and the absence of a performance gap is robust to using equal-weighted portfolios (bottom subtable) instead of value-weighted portfolios. This result is unsurprising as mutual funds are expected to be relatively good at implementing the market factor.

Broadening our focus to columns 1–4, we see that mutual funds under-perform stocks in isolating factor exposures for two of the other Carhart factors. The average implementation gaps for value (*HML*) and momentum (*UMD*) range from of 50% to 80% of the total on-paper factor return in stock portfolios. The remaining compensation to mutual funds for *HML* and *UMD* is positive ($\lambda^{MF} > 0$), but they are only 1%–3% per year and not statistically distinguishable from zero. Conversely, *HML* and *UMD* factors are both highly compensated and statistically robust in value-weighted stock portfolios in this period. On-paper compensation for size factor (*SMB*) exposure has a smaller positive point estimate, but this value is not reliably different from zero.

The point estimates for the differences λ^{Δ} for *HML* and *UMD* are typically more statistically significant than either of the components of the difference λ^S or λ^{MF}. This feature reflects the netting out of common variation in factor realizations between the λ time series. Ideally, the residual variation in λ^{Δ} captures only random variation in trading

[18] We verify in Appendix B that both stock– and mutual fund–based factor-mimicking portfolio returns replicate factor dynamics well. We thank Andrea Frazzini for this suggestion.

[19] See Appendix F for additional discussion.

Table 2
Implementation cost estimates, baseline specification.
This table reports Fama-MacBeth estimates of the compensation for factor exposure for stock portfolios (third and fourth rows), domestic equity mutual funds (fifth row), and their difference (first and second rows). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$,

$$r_{it} = \sum_k \lambda^S_{kt} \hat{\beta}_{ik} 1_{i \in S} + \sum_k \lambda^{MF}_{kt} \hat{\beta}_{ik} 1_{i \in MF} + \epsilon_{it}, \ t = 1, \ldots, T,$$

where k indexes the four Carhart (1997) factors and λ^Δ is defined as $\lambda^S - \lambda^{MF}$. Stock portfolio sets are described in Section 3. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1970 – 2016				1993 – 2016			
	N_S	*MKT*	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
Panel A: Value-weighted stock portfolios									
λ^Δ	100	−0.38	3.81***	0.26	7.18***	−0.11	3.12***	−0.24	4.27***
		(−1.28)	(5.08)	(0.42)	(5.53)	(−0.32)	(3.83)	(−0.29)	(2.64)
λ^Δ	269	−0.21	2.59***	−0.07	7.30***	0.28	2.09***	−0.97	5.04***
		(−0.88)	(3.81)	(−0.14)	(5.54)	(1.25)	(3.31)	(−1.39)	(2.89)
λ^S	100	6.60***	6.43***	1.27	8.72***	7.67**	5.43*	1.96	6.01
		(2.75)	(3.51)	(0.75)	(3.74)	(2.35)	(1.93)	(0.81)	(1.60)
λ^S	269	6.77***	5.20***	0.94	8.85***	8.06**	4.40	1.23	6.78*
		(2.82)	(2.84)	(0.56)	(3.80)	(2.49)	(1.54)	(0.51)	(1.83)
λ^{MF}	–	6.98***	2.62	1.01	1.54	7.78**	2.31	2.20	1.73
		(2.86)	(1.51)	(0.59)	(0.63)	(2.38)	(0.83)	(0.92)	(0.45)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123
Panel B: Equal-weighted stock portfolios									
λ^Δ	100	−0.36	4.47***	2.34**	6.83***	0.07	3.16***	2.14	3.71**
		(−0.76)	(5.57)	(2.41)	(5.21)	(0.12)	(3.29)	(1.55)	(2.21)
λ^Δ	269	0.25	3.31***	2.22**	8.51***	0.95	2.01*	2.05	6.04***
		(0.5)	(3.58)	(2.05)	(6.19)	(1.45)	(1.96)	(1.34)	(3.13)
λ^S	100	6.62***	7.09***	3.35***	8.37***	7.85**	5.48**	4.34	5.45
		(2.75)	(3.91)	(1.70)	(3.59)	(2.39)	(1.99)	(1.53)	(1.44)
λ^S	269	7.23***	5.93***	3.23	10.06***	8.73***	4.33	4.25	7.78**
		(3.02)	(3.03)	(1.56)	(4.17)	(2.69)	(1.47)	(1.43)	(1.98)
λ^{MF}	–	6.98***	2.62	1.01	1.54	7.78**	2.31	2.20	1.73
		(2.86)	(1.51)	(0.59)	(0.63)	(2.38)	(0.83)	(0.92)	(0.45)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123

costs. In practice, this residual variation also captures idiosyncratic differences in estimated risk prices associated with using different sets of test assets; the difference between λ^Δ estimated from the set of 100 stock portfolios and the set of 269 stock portfolios suggests that the implementation gap depends in part on the stock benchmarks employed.

Columns 5–8 reproduce these tests for the July 1993 to December 2016 sample. Mutual funds achieve lower returns to *HML* and *UMD* and higher returns to *SMB* than in the full sample, and these returns are universally statistically indistinguishable from zero. For stock portfolios, the compensation for *HML* and *UMD* (*SMB*) exposures also decreases (increases) relative to the full sample. The net effect of these changes is a small decrease in the typical implementation gap for *HML* and a moderate decrease in the implementation gap for *UMD*. The implementation gap is roughly unchanged for market exposure (effectively zero) and *SMB* exposure (positive but now statistically insignificant). Focusing on this latter subsample with a more broadly representative set of mutual funds does not change our conclusions on the high real-world efficacy of achieving market exposure and size and the low real-world efficacy of implementing value and momentum.

In sum, no factor other than the market earns reliably positive risk premia for the typical mutual fund. This finding is our first main result, and below we investigate whether it survives alternative weightings (including equal-weighted returns in the lower panel), enriched methodologies, additional controls, and sample splits.[20] Section 4.5 presents a detailed comparison of our estimates to prior work, and we present discussion of economic mechanisms and implications there. However, we

[20] For example, Section 4.4 evaluates a purely characteristic-based variant of our Fama-MacBeth regressions with similar results. The online supplemental appendix introduces a complementary, matched-pairs approach to investigating mutual fund implementation costs. In the spirit of Daniel et al. (1997), this approach compares returns to high book-to-market ratio, small size, and high prior return stocks and mutual funds with similar risk characteristics. The analysis therein has the ancillary benefit of controlling for differences in the distribution of betas between stock portfolios and mutual funds, which may be important if compensation for factor exposure is earned only in some segments of this distribution.

understand that implementation cost estimates of this magnitude may surprise some readers, and it is worth mentioning now an intuitive channel by which real-world compensation to these factors may fall to zero even as on-paper returns persist.

Using evidence from mutual fund flows, Berk and van Binsbergen (2016) and Barber et al. (2016) demonstrate that investors appear to use the CAPM to evaluate market risks, and compensation accruing to non-market sources is perceived as skill. By this logic, capital market equilibrium requires that real-world factor compensation must be squeezed to zero because these other factors are not seen as risky. Otherwise, if funds were to achieve positive returns to *HML* and *UMD*, funds could load on these factors to achieve alpha and attract inflows indefinitely. Our findings indicate that equilibrium obtains through greater implementation costs rather than through reduced on-paper factor compensation. Higher implementation costs resulting from greater fund size (Berk and Green, 2004) or industry scale (Pastor and Stambaugh, 2012) are two established mechanisms by which these costs can adjust until achievable anomalous returns disappear.

4.3. Estimates when costs vary across funds and time

Time-varying implementation costs complicate the comparison of compensation per unit of factor risk. To see why, consider the following augmented model of mutual fund costs. As above, let there be a set of academic factors f, where f_t is a $1 \times K$ vector. Each mutual fund i implements its favored version of academic factors and earns a return of

$$h_{it} = f_t - \eta_{it}, \tag{3}$$

where η_{it} reflects tilts away from the academic factor on account of trading costs or factor optimization. This section differs from Section 4.2 in that we no longer assume that η is constant across funds and time in interpreting λ^{Δ}. Allowing for cross-fund heterogeneity is particularly important in light of prior work by Edelen et al. (2007), Anand et al. (2011), and Edelen et al. (2013) showing substantial heterogeneity in execution and trading costs among mutual funds.

The η_{it} term in turn can be decomposed into components,

$$\eta_{it} = \eta_i + \eta_t \gamma_i + \tilde{\eta}_{it}. \tag{4}$$

The first component is the fixed, fund-specific cost of trading a factor. The second component is a set of L time-varying liquidity costs, η_t multiplied by the $L \times K$ loadings of all factors on these liquidity costs, γ_i. Finally, $\tilde{\eta}_{it}$ is a $1 \times K$ set of idiosyncratic costs, e.g., a surprise redemption demand that makes continued investment in factor k more costly for fund i.

In this heterogeneous cost specification, funds earn returns of

$$\begin{aligned} r_{it} &= \alpha_i + h_{it}\beta_i + \epsilon_{it} \\ &= (\alpha_i - \eta_i\beta_i) + (f_t - \eta_t\gamma_i)\beta_i + (\epsilon_{it} - \tilde{\eta}_{it}\beta_i). \end{aligned} \tag{5}$$

An ideal test compares the average compensation f_t for factor exposure for on-paper investment in stocks against the compensation h_{it} for factor exposure for real-world investment. In the constant cost setting of Section 4.1, we achieve this ideal: η_{it} simplifies to η, and Fama-MacBeth regressions recover consistent estimates of h as the difference in λs in Eq. (2).

In this general setting we face two key challenges that complicate the comparison of f_t and h_{it}. First, trading costs vary over time, and these costs can co-vary with factor realizations. For example, during the 2007–2008 financial crisis, the aggregate market declines sharply just as funding and market liquidity deteriorate significantly. Omitting relevant liquidity factors thus contributes to an omitted variable bias in time series estimates of β_i for investment managers, which in turn potentially invalidates simple comparisons of second-stage slope estimates. Second, investment managers select their risk exposures endogenously. An investor who has discovered improvements upon academic factors and another who faces particularly high trading costs are unlikely to select the same factor exposures, all else equal. For this reason, we would expect mutual fund-specific trading costs η_i to be correlated with β_i in the cross-section.

We now address these two sources of bias. First, to address the omission of trading cost factors, we assume that trading costs or optimization gains for mutual funds are spanned by liquidity proxies considered in the literature (and described in Section 3). To avoid overfitting by including too many correlated liquidity proxies, we start with two.[21] We use the first principal component of four market-liquidity variables [Amihud (2002) illiquidity, Pastor and Stambaugh (2003) liquidity, Corwin and Schultz (2012) bid-ask spreads, and the CBOE VIX/VXO] and the first principal component of four funding-liquidity variables [the Frazzini and Pedersen (2014) "betting against beta" factor, the He et al. (2017) intermediary capital risk factor, 10-year BAA minus 10-year Treasury spreads, and 3-month LIBOR minus 3-month Treasury yield or "TED" spreads].[22] We normalize all liquidity variables to have unit standard deviation before taking principal components because liquidity proxies vary widely in their scales. We assign these components an illiquidity interpretation by normalizing them to be positively correlated with the VIX/VXO.

We then run Fama-MacBeth regressions as before, but we extend the factor model to include these liquidity prox-

[21] Ideally we would use all liquidity variables instead of their principal components because we want time-varying determinants of η_{it} to lie in the span of the liquidity-augmented factor model. To this end, we include all proxies in a sparse-regression approach in Appendix D.

[22] The CBOE VXO and the TED spread series start in January 1986. Our principal components procedure accommodates the missing liquidity proxy data using MATLAB's alternating least squares (PCA-ALS) algorithm [based on Roweis (1998) and similar to the expectation-maximization procedure described in Appendix A of Stock and Watson (2002)]. PCA-ALS extracts factors and completes missing data by conjecturing principal components and iteratively estimating principal component loadings ϕ and factor values g until the distance between known and fitted values achieves a local minimum. We run PCA-ALS from one thousand starting points and select the global distance-minimizing factors and loadings.

ies in the time series regressions,

$$r_{it} = \alpha_i + \sum_k f_{kt}\beta_{ik} + \sum_l \tilde{\eta}_{lt}\tilde{\gamma}_{il} + \epsilon_{it},$$
$$i = 1, \ldots, N_S, N_{S+1}, \ldots, N_S + N_{MF}, \quad (6)$$

where $\tilde{\eta}_{lt}$ are the liquidity factor proxies at time t. The second-stage cross-sectional regressions are exactly as in Eq. (2).

The mismatch in model specification for the time series and cross-sectional regressions is intentional, and the decomposition of the resulting second-stage coefficient estimates reveals why the second source of bias, cross-sectional heterogeneity in implementation costs, makes our results conservative. In the time series regressions, we recover fund exposures to the academic factors, and we need the additional liquidity proxy variables to cleanse the estimated mutual fund factor loadings of omitted illiquidity components. By contrast, in the second stage, we recover the cost per unit exposure to the academic factors and do not want to include the liquidity proxy exposures. Excluding the liquidity factors only in the second stage delivers $\hat{\lambda}_t^S = \lambda_t^S$ and

$$\hat{\lambda}_t^{\Delta} = \lambda_t^S - \frac{cov\left(r_{it}^{MF}, \beta_i\right)}{var(\beta_i)} = -\frac{cov(\alpha_i - \eta_{it}\beta_i, \beta_i)}{var(\beta_i)}$$
$$= \bar{\eta}_t - \frac{cov((\bar{\eta}_t - \eta_{it})\beta_i, \beta_i)}{var(\beta_i)}. \quad (7)$$

The final equality makes the standard assumption that alphas and betas are cross-sectionally uncorrelated. $\bar{\eta}_t$ represents the cross-sectional average per-unit liquidity costs to implementing the factor. The second term is the covariance between deviations from the average costs and βs. Funds with a particular skill in investing in a factor likely have higher exposures to it, β_i is endogenous, so β_i is high when $\bar{\eta}_t - \eta_{it}$ is high, and β_i is close to zero when $\bar{\eta}_t - \eta_{it}$ is negative (negative betas do not reverse the sign on costs). Combining these features, the overall covariance is positive, and the cross-sectional slopes of returns with respect to β_i are biased upward ($\hat{\lambda}_t^{MF} > \lambda_t^{MF}$).[23] Consequently, λ_{kt}^{MF} is an upper bound on the realizable gains to factor investing per unit risk exposure, and λ_{kt}^{Δ} is a lower bound on the costs of implementing a factor strategy.

Table 3 presents results from the liquidity-extended first-stage regression. Results are virtually the same as those of the baseline specification in Table 2 with one exception. Mutual funds' (already low) annual compensation for *UMD* exposure decreases from 1.54% to 1.28% in the long sample and from 1.73% to 0.76% in the recent sample, suggesting that liquidity factor exposure at least partly explains mutual funds' compensation for momentum. Asness et al. (2013) find that momentum loads positively on liquidity risk, and we find that the same holds for mutual funds' implementation of momentum. We examine this feature in detail in Section 6.2.

[23] Including liquidity proxies in the second stage introduces a more opaque omitted variable bias, as we discuss in Appendix C.

4.4. Cross-sectional characteristic regressions

The Fama-MacBeth regression approach of the preceding sections estimates implementation costs for asset pricing factors under the assumption that factor exposures are the source of risk premia. However, Daniel and Titman (1997) and Daniel et al. (1997), among others, argue that characteristics such as book-to-market ratios and market capitalization dominate factors in explaining the cross-section of expected stock returns and mutual fund performance.[24] To address this class of models, we modify our baseline two-stage regression approach to use characteristics instead of factor betas. The resulting cross-sectional slopes are estimates of the compensation to characteristics accruing to on-paper stock portfolios and in real-world mutual funds.

We obtain characteristic prices in the style of Fama-MacBeth regressions by replacing the time series beta estimates from Eq. (1) with stock portfolio or fund characteristics, c_{ikt}, in the cross-sectional regressions,

$$r_{it} = \sum_k \lambda_{kt}^S c_{ikt} 1_{i\in S} + \sum_k \lambda_{kt}^{MF} c_{ikt} 1_{i\in MF} + \epsilon_{it}, \quad t = 1, \ldots, T. \quad (8)$$

Because characteristics are directly observed, not estimated, we no longer face an errors-in-variables problem arising from using estimated betas in the second-stage regression.[25] This feature allows us to use time-varying characteristics c_{ikt} rather than averages over the full time series. Indeed this replacement is critical because while betas are relatively stable across the sample period, characteristics such as average market capitalization have strong time trends.[26]

With the modified methodology in hand, the next step is to specify the set of characteristics and their construction. We follow Daniel and Titman (1997) and Daniel et al. (1997) in using market capitalization, book-to-market ratios, and prior returns as characteristics. We construct these characteristics at the stock-month level using book-to-market ratios from the most recent fiscal year, market capitalization at the end of the current month, and prior 12-month minus 2-month returns. We then lag book-to-market ratios and market capitalization by one month to ensure that all characteristics are available to market participants at the start of month t. To control data errors in the book-to-market ratio, we drop negative values and winsorize at the 1% level within each date.

[24] Berk (2000) and Davis et al. (2000) provide other views in the debate on compensation to factors or characteristics. We thank Juhani Linnainmaa and Ronnie Sadka for the suggestion to consider both perspectives.

[25] For this reason, characteristics are sometimes used in place of otherwise noisily estimated betas, or as instruments for them (e.g., Shanken, 1992; Kelly et al., 2018; Fama and French, 2018). Thus, the characteristics of Eq. (8) can be seen as a substitute for (imprecise) time-varying betas in evaluating mutual fund implementation costs.

[26] Because of these time trends and because a zero value of a characteristic need not command a zero risk premium (unlike betas), we also include a constant as part of the characteristic set. The inclusion of a constant at each date eliminates the influence of time trends by absorbing shifts in the means of the characteristics. By the same token, absorbing time-varying means of the characteristics renders the constant term uninterpretable, and we do not report it in our results.

Table 3
Implementation cost estimates including liquidity principal components.
This table reports Fama-MacBeth estimates of the compensation for factor exposure for stock portfolios (third and fourth rows), domestic equity mutual funds (fifth row), and their difference (first two rows). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$,

$$r_{it} = \sum_k \lambda^S_{kt} \hat{\beta}_{ik} 1_{i \in S} + \sum_k \lambda^{MF}_{kt} \hat{\beta}_{ik} 1_{i \in MF} + \epsilon_{it}, \quad t = 1, \ldots, T,$$

where k indexes the four Carhart (1997) factors and λ^Δ is defined as $\lambda^S - \lambda^{MF}$. First-stage regression estimates include these factors, the first principal component of market liquidity proxies, and the first principal component of funding liquidity proxies. Liquidity proxies and stock portfolio sets are described in Section 3. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1970 – 2016				1993 – 2016			
	N_S	*MKT*	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
Panel A: Value-weighted stock portfolios									
λ^Δ	100	−0.44	4.07***	0.35	7.49***	−0.12	3.30***	−0.24	5.23***
		(−1.45)	(5.17)	(0.57)	(5.71)	(−0.36)	(3.92)	(−0.29)	(3.09)
λ^Δ	269	−0.22	2.83***	−0.02	7.55***	0.27	2.30***	−0.93	5.84***
		(−0.92)	(3.87)	(−0.03)	(5.70)	(1.22)	(3.54)	(−1.32)	(3.29)
λ^S	100	6.55***	6.71***	1.26	8.77***	7.68**	5.38*	1.98	5.99
		(2.74)	(3.63)	(0.74)	(3.76)	(2.37)	(1.90)	(0.82)	(1.59)
λ^S	269	6.77***	5.47***	0.89	8.84***	8.08**	4.39	1.30	6.60*
		(2.83)	(2.94)	(0.53)	(3.78)	(2.51)	(1.51)	(0.54)	(1.78)
λ^{MF}	–	6.99***	2.64	0.90	1.28	7.80**	2.09	2.22	0.76
		(2.87)	(1.51)	(0.53)	(0.52)	(2.41)	(0.74)	(0.92)	(0.20)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123
Panel B: Equal-weighted stock portfolios									
λ^Δ	100	−0.53	4.48***	2.69***	6.92***	0.00	2.67***	2.32*	4.00**
		(−1.12)	(5.35)	(2.75)	(5.14)	(0.00)	(2.59)	(1.71)	(2.27)
λ^Δ	269	0.05	3.64***	2.64**	8.54***	0.74	1.92*	2.29	5.73***
		(0.09)	(3.79)	(2.44)	(5.95)	(1.09)	(1.81)	(1.56)	(2.82)
λ^S	100	6.46***	7.12***	3.60*	8.20***	7.81**	4.76*	4.54	4.76
		(2.70)	(3.88)	(1.84)	(3.51)	(2.40)	(1.73)	(1.63)	(1.26)
λ^S	269	7.04***	6.28***	3.55*	9.82***	8.55***	4.01	4.51	6.49
		(2.97)	(3.14)	(1.73)	(4.05)	(2.66)	(1.35)	(1.57)	(1.64)
λ^{MF}	–	6.99***	2.64	0.90	1.28	7.80**	2.09	2.22	0.76
		(2.87)	(1.51)	(0.53)	(0.52)	(2.41)	(0.74)	(0.92)	(0.20)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123

We then build characteristics at the stock portfolio and mutual fund group levels as value-weighted averages of the characteristics of their constituent stocks. For stock portfolios, we use Ken French breakpoints where available to partition NYSE, AMEX, and NASDAQ common stocks (share code 10 or 11). For illiquidity-sorted portfolios, we use quintiles of Amihud illiquidity in univariate-sorted portfolios and conditional quintiles of Amihud illiquidity by market capitalization bin in double-sorted portfolios. The stock portfolio value of each characteristic is the value-weighted average of its constituent stocks' characteristics.

To construct mutual fund characteristics, we first obtain mutual fund holdings using the Thomson Reuters mutual funds holdings database (s12). We match holdings at the fund level using MFLINKS to convert Thomson Reuters identifiers to CRSP mutual fund identifiers.[27] We form fund-level characteristics as the dollar holdings-weighted average of stock-level characteristics and fund-group characteristics as the TNA-weighted average of fund-level characteristics. Finally, we take logs of book-to-market ratios and market capitalization to prevent the regressions from being dominated by outlier firms.[28]

Table 4 reports results of Fama-MacBeth style regressions using our characteristic pricing model. By contrast with Daniel and Titman (1997), we do not find strong evidence of compensation for characteristics in stock or mutual fund portfolios. This result is likely due to the sensi-

[27] Details on the merge procedure are available at the Guide for MFLINKS on Wharton Data Research Services (WRDS); most importantly for our application, the link table matches up to 98% of the domestic equity funds in CRSP for the March 1980 to September 2015 period in which linking data are available.

[28] Characteristic ranks are an alternative transformation sometimes used in characteristic regressions, but they are inappropriate in our setting for two reasons. First, a change in rank has a different meaning for stock portfolios and mutual funds, particularly in light of the flow-performance relation related to mutual funds' prior returns characteristic. Second, the distribution of characteristics differs for stock portfolios and mutual funds, so ranking must be performed across all entities so as to not destroy information about differences in average characteristics. However, doing so introduces the undesirable feature that stock portfolio characteristics depend on the set of mutual funds considered and vice versa.

Table 4
Implementation cost estimates using a characteristic model.
This table reports Fama-MacBeth estimates of the compensation for characteristic exposure for stock portfolios (third and fourth rows), domestic equity mutual funds (fifth row), and their difference (first two rows). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on characteristics,

$$r_{it} = \left(\lambda^S_{0t} + \lambda^S_{Bt} BM_{it} + \lambda^S_{St} SIZE_{it} + \lambda^S_{Pt} P212_{it}\right) 1_{i\in S} + \left(\lambda^{MF}_{0t} + \lambda^{MF}_{Bt} BM_{it} + \lambda^{MF}_{St} SIZE_{it} + \lambda^{MF}_{Pt} P212_{it}\right) 1_{i\in MF} + \epsilon_{it}, \ t = 1, \ldots, T,$$

where BM denotes lagged log book-to-market ratios, $SIZE$ denotes lagged log market capitalization, and $P212$ denotes prior 2–12 month return. λ^Δ is defined as $\lambda^S - \lambda^{MF}$. Stock portfolio sets are described in Section 3. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1980 – 2015			1993 – 2015		
	N_S	*BM*	*SIZE*	*P212*	*BM*	*SIZE*	*P212*
Panel A: Value-weighted stock portfolios							
λ^Δ	100	1.86	−0.37	10.64*	2.10	−0.51	9.04
		(1.07)	(−0.97)	(1.90)	(0.89)	(−1.10)	(1.25)
λ^Δ	269	−0.06	−0.11	7.16*	−1.33	−0.14	7.99*
		(−0.04)	(−0.42)	(1.92)	(−0.78)	(−0.39)	(1.69)
λ^S	100	2.53	−0.69	16.07**	2.17	−0.98	10.55
		(1.28)	(−1.21)	(2.36)	(0.82)	(−1.38)	(1.14)
λ^S	269	0.67	−0.43	13.02**	−1.14	−0.60	10.17
		(0.58)	(−1.07)	(2.49)	(−0.76)	(−1.16)	(1.41)
λ^{MF}	–	0.69	−0.34	5.70*	0.10	−0.49	1.91
		(0.48)	(−0.90)	(1.82)	(0.05)	(−1.06)	(0.44)
T		429	429	429	267	267	267
$\bar{N}_{MF}$		997	997	997	1405	1405	1405
Panel B: Equal-weighted stock portfolios							
λ^Δ	100	2.78	−0.44	10.48*	3.60	−0.62	10.39
		(1.52)	(−1.07)	(1.80)	(1.43)	(−1.24)	(1.36)
λ^Δ	269	−0.17	0.18	13.07***	−1.22	0.16	16.20***
		(−0.12)	(0.64)	(3.71)	(−0.70)	(0.42)	(3.53)
λ^S	100	3.35	−0.76	15.89**	3.51	−1.08	11.87
		(1.60)	(−1.36)	(2.29)	(1.24)	(−1.59)	(1.24)
λ^S	269	0.47	−0.14	18.82***	−1.19	−0.31	18.21***
		(0.47)	(−0.49)	(3.93)	(−0.90)	(−0.90)	(2.71)
λ^{MF}	–	0.69	−0.34	5.70*	0.10	−0.49	1.91
		(0.48)	(−0.90)	(1.82)	(0.05)	(−1.06)	(0.44)
T		429	429	429	267	267	267
$\bar{N}_{MF}$		997	997	997	1405	1405	1405

tivity of characteristic-based pricing models to the choice of functional form, and average return compensation may not be linear in logs. The prior returns characteristic is highly compensated in both value- and equal-weighted stock portfolios: a 1% increase in prior return is associated with a 10–19 basis point increase in future returns. Turning to mutual funds, compensation to this characteristic is a far lower 1.9–5.7 basis points, and only the latter value is even marginally statistically significant. Hence, the implementation gap on the momentum characteristic remains prohibitively high at 56%–90% of the on-paper stock portfolio compensation. Paralleling our results in Sections 4.2–4.3, we conclude that mutual funds cannot reliably earn premia on characteristic versions of any of the Carhart anomalies.

4.5. Comparison with cost estimates from other work

Table 5 compares our real-world factor return estimates with estimates from selected works in the literature. Novy-Marx and Velikov (2016) estimate trading costs by summing effective bid-ask spreads of traded securities, and by their reckoning, momentum's trading costs reduce the gross strategy return from 16.0% per year to 8.16% per year (Table 3 of their paper). These positive momentum returns net-of-costs likely significantly overstate achievable returns, however, because their calculation ignores the price impact of trading that is particularly relevant to institutional investors.

Papers that consider price impact costs reach mixed conclusions on the implementability of momentum. Korajczyk and Sadka (2004) suggest that momentum profits exist only at small scales (the table reports only their returns net of proportional costs, and, by their reckoning, non-proportional costs quickly overwhelm strategy returns), and Lesmond et al. (2004) argue that high transaction costs preclude profitable momentum strategies altogether. Because these studies estimate transactions-cost functions using all TAQ transactions, their average imple-

Table 5
Comparison with selected factor profitability estimates from prior work.
This table presents estimates of factor strategy returns. The first set of estimates are cross-sectional slopes using Fama-MacBeth regressions from Table 2. For brevity we report only the estimates in which liquidity proxy principal components appear in the time series step, and we focus on the slopes for the full sample of mutual funds and for small mutual funds (lagged total net assets between $10 million and $50 million). Standard errors are Newey-West with three lags. The second set of estimates are value-weighted momentum strategy returns from Table IV of Korajczyk and Sadka (2004). Alphas are constructed relative to the Fama-French three factors. $\alpha_{net}^{espr.}$ and $\alpha_{net}^{qspr.}$ represent excess momentum returns net of proportional costs as measured by effective spreads and quoted spreads, respectively. The third set of estimates are equal-weighted strategy returns from Table 3 of Lesmond et al. (2004) (value-weighted returns are not reported). r_{net}^{LDV} and r_{net}^{direct} are momentum returns net of Lesmond et al. (1999)-implied costs and "direct" costs (consisting of bid-ask spreads and trading commissions), respectively. The fourth set of estimates are realized strategy returns from Table IV of Frazzini et al. (2015). The final set of estimates are value-weighted strategy returns net of Hasbrouck (2009)–implied effective spreads from Table 3 of Novy-Marx and Velikov (2016). Throughout returns are annualized and *t* statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

Study (sample period)	Parameter	*HML*	*SMB*	*UMD*
This paper	λ^{MF}	2.64	0.90	1.28
(1970–2016)		(1.51)	(0.53)	(0.52)
	λ_{small}^{MF}	2.55	1.37	2.62
		(1.37)	(0.82)	(0.97)
Korajczyk and Sadka (2004)	α_{gross}			6.84***
(1967–1999)				(4.54)
	$\alpha_{net}^{espr.}$			5.40***
				(3.59)
	$\alpha_{net}^{qspr.}$			4.80***
				(3.17)
Lesmond et al. (2004)	r_{gross}			7.83***
(1980–1998)				(6.22)
	r_{net}^{LDV}			0.13
				(0.07)
	r_{net}^{direct}			2.24
				(1.22)
Frazzini et al. (2015)	r_{gross}	4.86	7.98***	2.26
(1986–2013)		(1.12)	(3.01)	(0.40)
	r_{net}	3.51	6.52**	-0.77
		(0.80)	(2.48)	(-0.14)
Novy-Marx and Velikov (2016)	r_{gross}	5.64***	3.96*	15.96***
(1963–2013)		(2.68)	(1.66)	(4.80)
	r_{net}	5.04**	3.36	8.16**
		(2.39)	(1.44)	(2.45)

mentation cost estimates smooth over heterogeneous investors and over trades unrelated to momentum strategies. Nevertheless, Lesmond et al. (2004) find that momentum has an economically unimportant premium for the average trader.

Our factor compensation estimates fall on the lower end of the spectrum, and our results are most similar to Lesmond et al. (2004) in that we find no net-of-cost compensation to momentum. We square our implementation cost estimates with prior work in two ways. First, we decompose implementation costs to better understand what frictions erode mutual funds' ability to capture factor premia. Section 5 considers the roles of shorting frictions and limitations on funds' investable universe, as well as the trade-off between tracking error and performance more generally. Second, Section 6 considers cross-sectional and time series variation in costs across funds, and we find substantial heterogeneity. Differences between average and skilled funds reconcile the lower costs seen in studies of single funds using proprietary trading data and studies of average traders in TAQ.

5. Decomposing implementation costs

5.1. The role of mutual fund shorting constraints

The implementation gap we estimate reflects the difference between on-paper and real-world performance for zero-cost factor strategies. Such strategies consist of financing long position by shorting other stocks, for example, selling growth stocks to purchase value stocks. Institutional impediments to shorting may significantly increase costs on the short side and reduce the performance of real-world factor strategies.

Table 6
Implementation cost estimates based on long-only factors.
This table reports Fama-MacBeth estimates of the compensation for long-only factor exposure for positive and negative beta stock portfolios (third and fourth rows), domestic equity mutual funds (fifth row), and their difference (first two rows). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}^{+}_{ik}$,

$$r_{it} = \sum_k \lambda^S_{kt} \hat{\beta}^{+}_{ik} 1_{i \in S} + \sum_k \lambda^{MF}_{kt} \hat{\beta}^{+}_{ik} 1_{i \in MF} + \epsilon_{it}, \ t = 1, \ldots, T,$$

where k indexes the long-only versions of the Carhart (1997) factors (the excess returns on *MKT*, *H*, *S*, and *U*) and λ^{Δ} is defined as $\lambda^S - \lambda^{MF}$. Stock portfolio sets are described in Section 3. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1970 – 2016				1993 – 2016			
	N_S	*MKT*	HML^+	SMB^+	UMD^+	*MKT*	HML^+	SMB^+	UMD^+
Panel A: Value-weighted stock portfolios									
λ^{Δ}	100	−0.61*	2.56***	0.52	3.09***	−0.32	2.10***	0.32	2.41***
		(−1.94)	(4.05)	(1.00)	(4.52)	(−0.92)	(3.85)	(0.56)	(2.98)
λ^{Δ}	269	−0.29	1.60***	0.02	2.85***	0.18	1.30***	−0.34	2.46***
		(−1.21)	(2.72)	(0.04)	(4.25)	(0.81)	(3.00)	(−0.61)	(2.85)
λ^S	100	6.22***	12.25***	9.19***	11.69***	7.32**	12.89***	10.84**	12.12***
		(2.59)	(4.33)	(2.85)	(4.11)	(2.24)	(3.19)	(2.54)	(3.22)
λ^S	269	6.54***	11.29***	8.68***	11.46***	7.82**	12.09***	10.17**	12.17***
		(2.73)	(3.95)	(2.68)	(4.02)	(2.41)	(2.95)	(2.38)	(3.24)
λ^{MF}	–	6.83***	9.69***	8.66***	8.60***	7.63**	10.80***	10.51**	9.71**
		(2.81)	(3.25)	(2.60)	(2.85)	(2.34)	(2.59)	(2.44)	(2.48)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123
Panel B: Equal-weighted stock portfolios									
λ^{Δ}	100	−0.83	3.82***	2.21***	3.64***	−0.25	3.11***	2.31***	3.27***
		(−1.66)	(6.06)	(3.37)	(5.46)	(−0.43)	(5.57)	(2.64)	(4.01)
λ^{Δ}	269	−0.26	3.40***	2.17***	4.25***	0.41	2.84***	2.38**	4.26***
		(−0.50)	(4.92)	(2.96)	(6.29)	(0.66)	(4.21)	(2.25)	(4.86)
λ^S	100	6.00**	13.50***	10.88***	12.24***	7.38**	13.91***	12.82***	12.98***
		(2.49)	(4.71)	(3.27)	(4.27)	(2.23)	(3.41)	(2.87)	(3.39)
λ^S	269	6.57***	13.09***	10.83***	12.86***	8.04**	13.64***	12.90***	13.98***
		(2.75)	(4.44)	(3.19)	(4.44)	(2.46)	(3.23)	(2.80)	(3.64)
λ^{MF}	–	6.83***	9.69***	8.66***	8.60***	7.63**	10.80***	10.51**	9.71**
		(2.81)	(3.25)	(2.60)	(2.85)	(2.34)	(2.59)	(2.44)	(2.48)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123

In this section we adapt our Fama-MacBeth approach to evaluate the extent to which implementation gaps arise from shorting frictions. To do this, we consider two long-only variants of value, size, and momentum. For pure no-shorting strategies, mutual funds borrow at the risk-free rate to invest in the long side of each factor. Our long-only factors are the excess returns on *H*, *S*, and *U* portfolios, all of which are accessible to short-sale constrained mutual funds, and we denote these long-only factors with a "+" superscript.

The typical mutual fund is highly exposed to the market, the mean and median correlations with the S&P 500 are 85% and 90%, and increasing exposure to *H*, *S*, or *U* can be financed by reducing a long position in other securities (e.g., the market) instead of by opening a short position. With this motivation, we also consider returns on "tilt" factors, defined as the difference between the long-factor portfolios and the market.[29] We denote the tilt factors with a "#" superscript. For both sets of factors, we do not modify *MKT* because the market factor is already in excess return form and accessible to long-only funds.

Table 6 reports stock portfolio and mutual fund returns to the long-only Carhart factors. Focusing on the differences in premia earned, λ^{Δ}, relative to the baseline estimates, the long-only factor implementation costs are about 60% as large for HML^+ and about 40% as large for UMD^+, but they are of comparable statistical significance.[30] As before, we find no evidence of significant implementation costs for market or long-only size factor exposures in value-weighted portfolios. Equal-weighted portfolio results

[29] Such tilt factors also have the advantage of closely tracking the traditional Carhart factors. For example, if *H* and *L* have comparable market capitalization for all dates, then the return to the tilt factor $HML^{\#}$ is $H - (H+L)/2$ or $HML/2$, and the accessible tilt factor is proportional to standard *HML*.

[30] Intriguingly, the scaling of long-only implementation costs relative to total implementation costs is in line with the Israel and Moskowitz (2013) finding that roughly 60% of the value premium and 50% of the momentum premium are earned on the long side of the anomalies. We find that real-world trading costs are roughly proportional to premia earned in on-paper portfolios.

Table 7
Implementation cost estimates based on tilt factors.
This table reports Fama-MacBeth estimates of the compensation for long-only factor exposure for positive and negative beta stock portfolios (third and fourth rows), domestic equity mutual funds (fifth row), and their difference (first two rows). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}^{\#}_{ik}$,

$$r_{it} = \sum_k \lambda^S_{kt} \hat{\beta}^{\#}_{ik} 1_{i \in S} + \sum_k \lambda^{MF}_{kt} \hat{\beta}^{\#}_{ik} 1_{i \in MF} + \epsilon_{it}, \ \ t = 1, \ldots, T,$$

where k indexes "tilt" versions of the Carhart (1997) factors (the excess return on the market, and H, S, and U net of the market) and λ^{Δ} is defined as $\lambda^S - \lambda^{MF}$. Stock portfolio sets are described in Section 3. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1970 – 2016				1993 – 2016			
	N_S	*MKT*	*HML*#	*SMB*#	*UMD*#	*MKT*	*HML*#	*SMB*#	*UMD*#
Panel A: Value-weighted stock portfolios									
λ^{Δ}	100	−0.61*	3.17***	1.13*	3.70***	−0.32	2.41***	0.64	2.72***
		(−1.94)	(4.34)	(1.77)	(5.08)	(−0.92)	(3.39)	(0.86)	(3.09)
λ^{Δ}	269	−0.29	1.89***	0.31	3.15***	0.18	1.11**	−0.52	2.27***
		(−1.21)	(3.06)	(0.58)	(4.81)	(0.81)	(2.36)	(−0.91)	(2.74)
λ^S	100	6.22***	6.03***	2.97*	5.47***	7.32**	5.58**	3.52	4.80***
		(2.59)	(4.19)	(1.94)	(4.57)	(2.24)	(2.49)	(1.64)	(2.72)
λ^S	269	6.54***	4.75***	2.15	4.92***	7.82**	4.28*	2.36	4.35**
		(2.73)	(3.28)	(1.42)	(4.18)	(2.41)	(1.85)	(1.11)	(2.54)
λ^{MF}	–	6.83***	2.86*	1.83	1.77	7.63*	3.16	2.88	2.08
		(2.81)	(1.92)	(1.17)	(1.47)	(2.34)	(1.36)	(1.34)	(1.24)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123
Panel B: Equal-weighted stock portfolios									
λ^{Δ}	100	−0.83	4.65***	3.04***	4.47***	−0.25	3.36***	2.56**	3.52***
		(−1.66)	(5.56)	(3.16)	(5.92)	(−0.43)	(4.31)	(2.18)	(3.81)
λ^{Δ}	269	−0.26	3.66***	2.43**	4.52***	0.41	2.43**	1.97	3.85***
		(−0.50)	(3.97)	(2.28)	(6.33)	(0.66)	(2.55)	(1.38)	(4.34)
λ^S	100	6.00**	7.50***	4.88***	6.24***	7.38**	6.53***	5.44**	5.60***
		(2.49)	(5.06)	(2.75)	(4.98)	(2.23)	(2.98)	(2.23)	(3.08)
λ^S	269	6.57***	6.52***	4.27**	6.29***	8.04**	5.60**	4.85*	5.93***
		(2.75)	(4.00)	(2.26)	(4.82)	(2.46)	(2.27)	(1.82)	(3.23)
λ^{MF}	–	6.83***	2.86*	1.83	1.77	7.63*	3.16	2.88	2.08
		(2.81)	(1.92)	(1.17)	(1.47)	(2.34)	(1.36)	(1.34)	(1.24)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123

are very similar, although we do find a significant SMB^+ implementation gap because of the increased weight assigned to difficult-to-access microcaps with high average returns.

The table also reports the risk premia earned on the long side of each factor. On paper, long-only value and momentum premia are large relative to the equity premium, with gaps of 5%–6% depending on the choice of portfolio weighting and time period. These expected-return improvements earned by tilting away from the market portfolio are reflected much less strongly in mutual funds.

Table 7 replaces long-only factors with tilt factors. Our conclusions are much the same as above: $HML^{\#}$ and $UMD^{\#}$ suffer large implementation costs in practice regardless of time period or portfolio weighting. Moreover, the magnitude of these estimated costs is comparable to that of the long-only factors in Table 6: real-world under-performance is robust to assumptions on how funds implement the long side of anomalies. Focusing on the second and third panels delivers statistical assessments of the cross-column comparisons of Table 6. Mutual funds earn a marginally statistically significant premium on value tilts in the full sample and zero premium to factor tilts for all other factors and sample periods. By contrast, stocks earn robust premia to value and momentum tilts.

From both tables we conclude that the implementation costs of long-only versions of standard factors are significant and comparable to short-side costs. Real-world shorting frictions hence explain as much as half of the high all-in implementation costs of value and momentum factors. The high cost of shorting restrictions may explain the growing popularity of levered mutual funds, e.g., "130/30" funds, for which these restrictions are less binding.

5.2. The role of investability frictions

Implementation costs attenuate the returns to traded securities and motivate investors to depart from prescribed factor strategies. Frictions that reduce the set of investment opportunities are an important "shadow" implementation cost, analogous to the shadow price on a constraint on which stocks can be included in a portfolio, faced by real-world investors and missed by existing measures of costs. In this section we consider the role of security size in circumscribing mutual funds' investable universe. Security size is a natural candidate for explaining the performance gap between on-paper and real-world factor investing because the highest returns to *HML* exposure are earned in the smallest stocks (Fama and French, 2012; Israel and Moskowitz, 2013), and low-market capitalization securities are too small to accommodate meaningful investment by large mutual funds.

The smallest stocks, or "microcaps," present especially challenging environments for asset managers because of their particularly low carrying capacities and high transaction costs. Perhaps because of the challenges facing potential arbitrageurs in this space, the majority of academic anomalies exist only in these "dusty corners" of the stock market (Hou et al., 2017). To evaluate the effect of microcaps on our cost estimates, we exclude microcaps from our set of stock portfolios. We follow Fama and French (2008) and Hou et al. (2017) in defining microcaps as stocks with market capitalization less than the 20th percentile of NYSE market capitalization, and we implement this filter by dropping the smallest-size portfolios from double-sorted size-value, size-beta, size-prior return, and size-Amihud portfolios. This exclusion eliminates a fifth of the portfolios but only 3% of market capitalization (Fama and French, 2008).

Table 8 reports Fama-MacBeth estimates of factor premia on this set of stock portfolios. We present only value-weighted results because we are interested in downweighting tiny stocks to reflect the investable universe. Our main finding is that microcaps indeed explain some of the measured performance attrition for value and momentum strategies, but not enough to close the measured implementation gap. As a useful placebo, the gap on replicating performance on the value-weighted market changes by at most a few basis points.

In the 1970–2016 sample, both value and momentum compensation are about 1% smaller in the stock portfolios in which microcaps are excluded. This difference persists for value in the more recent sample, echoing Fama and French (2012) and Israel and Moskowitz (2013), but it roughly halves for momentum. Nevertheless, the performance gap between non-microcap stock portfolios and mutual funds remains economically large and statistically robust. If mutual funds indeed cannot invest in microcap stocks, this narrowing of the investable universe explains about one-third of the implementation gap for value and about one-sixth of the implementation gap for momentum.

5.3. Tracking error and the performance of factor strategies

Mutual funds face a trade-off between following high-cost canonical factor strategies and deviating from those strategies to capture the bulk of factor premia at lower costs. Benchmark-based performance evaluation in particular pushes funds to mimic factor benchmarks despite the potentially lower Sharpe ratios of doing (Basak and Pavlova, 2013). In this section we split our sample into quintiles by Carhart four-factor R^2s to evaluate whether variation in tracking error is associated with factor strategy performance.

This split also serves a second function in combating bias in our implementation cost estimates that arises from misspecifying mutual fund strategies. Bias occurs if incidental factor exposures incurred by other activities are cross-sectionally correlated with mutual fund returns. For especially high R^2 values in the time series regressions, the scope for omitted variable bias is small if coefficients are stable across specifications, as they are in our study (Oster, 2019).

To perform this split, we run the time series regressions fund-by-fund as before using the Carhart (1997) model, and we sort funds into one of five equally-spaced bins at each date based on the R^2 of their time series regression.[31] Funds with high R^2 have returns nearly spanned by the academic strategies, and these funds have low tracking error and little scope for omitted strategies that might complicate the interpretations of λ^{MF} and λ^{Δ}.[32] Conversely, funds with lower R^2 either implement academic strategies with greater discretion and/or tracking error, or implement strategies that we cannot observe. We then construct cross-sectional mutual fund factor compensation estimates for each R^2 group as in Tables 2–3.

Table 9 presents results from the splits by explanatory power of the four-factor model on the full 1970–2016 sample. The decomposition by R^2 delivers two results related to how funds implement asset pricing factors. First, perhaps unsurprisingly, funds that track factors more closely are generally more efficient at earning factor premia. Performance differences across fund quintiles are statistically significant for *MKT, SMB*, and *UMD* factors. This result is reversed for the market factor. Funds with greater deviations from the academic factors typically achieve greater returns to market beta. Small deviations from the CRSP market go a long way toward improving returns: market exposure is compensated 41–81 basis points more in mutual funds that track the four factors less well. This finding reinforces the importance of using flexible approaches to measuring implementation costs that are robust to real-world departures from the academic factors.

Second, and more importantly for our study, relative to the average mutual fund, funds with the highest R^2s achieve economically similar performance on market and value factors and somewhat higher performance on size and momentum factors. Value premia are about 1% larger among the funds most closely mimicking academic factors, and compensation for value exposure is significantly different from zero at the 5% level for the highest R^2 quintile. Likewise, returns to momentum exposure for this group are nearly triple those of the typical mutual fund, and they are statistically significant with strength depending

[31] Appendix G considers an alternative way to focus on funds that target particular factors. We find that funds with reliable exposures to a factor suffer only slightly lower costs than the typical mutual fund.

[32] The top R^2 quintile also includes many index funds (roughly 5% of observations in our sample). Because index funds seek to replicate factors at the lowest possible cost, we expect factor compensation estimates for this quintile to represent the best case achievable for passive or active mutual funds.

Table 8
Implementation cost estimates with microcaps excluded.
This table reports Fama-MacBeth estimates of the compensation for factor exposure in value-weighted stock portfolios in the baseline regressions (Panel A) and regressions with liquidity principal components (Panel B). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$,

$$r_{it} = \sum_k \lambda_{kt}^S \hat{\beta}_{ik} 1_{i\in S} + \sum_k \lambda_{kt}^{MF} \hat{\beta}_{ik} 1_{i\in MF} + \epsilon_{it}, \; t = 1, \ldots, T,$$

where k indexes the four Carhart (1997) factors and λ^Δ is defined as $\lambda^S - \lambda^{MF}$. First-stage regression estimates in Panel B include these factors, the first principal component of market liquidity proxies, and the first principal component of funding liquidity proxies. Liquidity proxies and stock portfolio sets are described in Section 3, with the important distinction that all portfolios with the smallest market capitalization quintile are excluded in the $N_S = 80$ specifications. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1970 – 2016				1993 – 2016			
	N_S	*MKT*	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
Panel A: Baseline specification									
λ^Δ	80	−0.37	2.85***	0.70	6.14***	0.10	2.06**	−0.42	3.49**
		(−1.39)	(3.84)	(1.24)	(4.66)	(0.32)	(2.44)	(−0.61)	(2.12)
λ^S	80	6.61***	5.47***	1.71	7.68***	7.88***	4.37	1.78	5.23
		(2.74)	(3.03)	(1.07)	(3.31)	(2.40)	(1.57)	(0.80)	(1.41)
λ^S	100	6.60***	6.43***	1.27	8.72***	7.67**	5.43*	1.96	6.01
		(2.75)	(3.51)	(0.75)	(3.74)	(2.35)	(1.93)	(0.81)	(1.60)
λ^{MF}	–	6.98***	2.62	1.01	1.54	7.78**	2.31	2.20	1.73
		(2.86)	(1.51)	(0.59)	(0.63)	(2.38)	(0.83)	(0.92)	(0.45)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123
Panel B: Including liquidity principal components									
λ^Δ	80	−0.37	2.97***	0.55	6.42***	0.13	2.25***	−0.69	4.55***
		(−1.34)	(3.81)	(0.97)	(4.82)	(0.4)	(2.64)	(−0.96)	(2.65)
λ^S	80	6.63***	5.60***	1.45	7.70***	7.93**	4.34	1.53	5.31
		(2.76)	(3.10)	(0.91)	(3.31)	(2.43)	(1.56)	(0.68)	(1.43)
λ^S	100	6.55***	6.71***	1.26	8.77***	7.68**	5.38*	1.98	5.99
		(2.74)	(3.63)	(0.74)	(3.76)	(2.37)	(1.90)	(0.82)	(1.59)
λ^{MF}	–	6.99***	2.64	0.90	1.28	7.80**	2.09	2.22	0.76
		(2.87)	(1.51)	(0.53)	(0.52)	(2.41)	(0.74)	(0.92)	(0.20)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123

on specification. Even so, the funds that most closely track the four academic factors continue to significantly underperform the on-paper factors. The best-performing R^2 segments for value see an implementation gap of 1%–2% relative to the stock portfolios, and the momentum implementation gap for these funds is roughly half the on-paper momentum premium.

One potential concern with sorting by R^2s is that doing so is equivalent to sorting on the precision in β estimates. Because attenuation toward zero increases in measurement error, compensation to factor premia may be biased downward more for low R^2 funds than high R^2 funds. In unreported results we apply the same Jegadeesh et al. (2019) measurement-error correction described in Appendix F. Results are unchanged throughout Table 9, with one notable exception: unlike the preceding results, the momentum implementation gap for the most successful segment is a third of the on-paper momentum premium in the full sample and a sixth in the recent sample. We leave a fuller analysis of cross-sectional differences in mutual funds' ability to harvest factor premia to ongoing work.

6. Cost estimates across funds and time

6.1. Implementation costs across funds

With the exception of the breakdown by four-factor R^2s, our analysis thus far considers the implementation costs of factor strategies for an average mutual fund, with no attention paid to heterogeneous characteristics and costs. Variation in investors' trading technologies may drive a wedge between a typical asset manager and the marginal investor in an anomaly. By dividing asset managers into groups we can learn whether factors are broadly (in)accessible or whether they generate positive net-of-costs returns for a subset of managers. In this section, we briefly demonstrate the utility of our cross-sectional approach for examining segments of asset managers.

Motivated by extensive work relating fund size to gross-of-fees performance (e.g., Berk and Green, 2004; Pastor et al., 2015; Berk and van Binsbergen, 2015), we split fund groups into groups based on lagged total net assets. We then run our second-stage cross-sectional regressions, presented in Eq. (2), separately for each asset manager TNA

Table 9
Fama-MacBeth slopes for stocks and mutual funds, R^2 quintile splits.
Table reports Fama-MacBeth estimates of the compensation for factor exposure for domestic equity mutual funds. Coefficients are the average cross-sectional slopes $\bar{\lambda}_k^g$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$ for each group of mutual funds g,

$$r_{it} = \sum_k \lambda_{kt}^{MF,g} \hat{\beta}_{ik} + \epsilon_{it}, \; t = 1, \ldots, T, \; g = 1, \ldots, 5,$$

where k indexes the four Carhart (1997) factors. We partition mutual funds into five equal groups sorted by time series regression R^2s from the Carhart model, where R^2 cutoffs are set at each date based on the sample of live funds. "5" indicates the highest R^2 funds, and "1" indicates the lowest R^2 funds. The first column reports the average R^2s across all fund-date observations within each R^2 group. Subsequent first-stage regression estimates include these factors only (columns two to five) and the first principal component of market and funding liquidity proxies (columns six to nine). Liquidity proxies and stock portfolio sets are described in Section 3. λ_5^Δ is the difference between compensation for factor exposure between the 269 stock portfolios and the highest R^2 fund group. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks. The final three rows report the p values of F tests of coefficients being jointly different from zero, of F tests of equality of coefficients, and the Patton and Timmermann (2010) test of non-monotonicity of coefficients.

		Baseline specification				Including liquidity PCs			
	$\bar{R}^2$	*MKT*	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
λ_5^{MF}	94.2%	6.50***	3.60**	1.78	4.59*	6.52***	3.97**	1.74	5.36**
		(2.69)	(1.99)	(1.04)	(1.68)	(2.71)	(2.15)	(1.02)	(1.99)
λ_4^{MF}	89.9%	6.91***	2.93*	2.67	0.73	6.89***	3.04*	2.72	0.34
		(2.82)	(1.70)	(1.57)	(0.26)	(2.83)	(1.74)	(1.59)	(0.12)
λ_3^{MF}	86.0%	7.31***	3.00*	0.09	3.23	7.28***	2.89	0.11	2.30
		(2.96)	(1.68)	(0.05)	(1.20)	(2.95)	(1.59)	(0.06)	(0.86)
λ_2^{MF}	79.9%	7.29***	2.66	1.15	−0.81	7.27***	2.42	1.14	−1.45
		(2.98)	(1.48)	(0.64)	(−0.31)	(2.98)	(1.32)	(0.64)	(−0.55)
λ_1^{MF}	55.4%	7.00***	2.93	−0.98	2.08	7.14***	3.44*	−1.51	2.38
		(2.80)	(1.52)	(−0.49)	(0.72)	(2.85)	(1.75)	(−0.75)	(0.81)
λ^{MF}	81.1%	6.98***	2.62	1.01	1.54	6.99***	2.64	0.90	1.28
		(2.86)	(1.51)	(0.59)	(0.63)	(2.87)	(1.51)	(0.53)	(0.52)
λ^S	–	6.77***	5.20***	0.94	8.85***	6.77***	5.47***	0.89	8.84***
		(2.82)	(2.84)	(0.56)	(3.80)	(2.83)	(2.94)	(0.53)	(3.78)
λ_5^Δ	–	0.27	1.60**	−0.84	4.26***	0.25	1.49*	−0.85	3.48**
		(1.17)	(2.02)	(−1.45)	(2.80)	(1.08)	(1.85)	(−1.46)	(2.21)
$\lambda_i = 0$		0.00***	0.41	0.00***	0.02**	0.00***	0.22	0.00***	0.00***
$\lambda_i = \lambda^*$		0.00***	0.83	0.00***	0.01***	0.01**	0.47	0.00***	0.00***
$\Delta\lambda \lessgtr 0$		0.27	0.08*	0.79	0.68	0.13	0.37	0.85	0.78

group.[33] We set aside funds with less than \$10 million in assets because selection into this group implies that the fund has lost money (we retain observations only after funds reach \$10 million in assets to avoid incubation bias).

Table 10 presents results from these segmented regressions on the full 1970–2016 sample. As in Tables 2–3, mutual funds generally achieve returns to market factor exposure comparable to those of on-paper stock portfolios. *HML* also earns positive compensation for most TNA groups, but returns to *HML* are not statistically different from zero in both specifications, with the possible exception of the mega-funds group. Point estimates for returns to *SMB* are positive for all fund size groups excluding micro funds, but *SMB* compensation estimates are not statistically distinguishable from zero or from each other.

Focusing on momentum, we estimate large differences in compensation across mutual fund size categories, with the smallest funds earning 5%–6% more per unit momentum beta than the largest funds. Notwithstanding the greater momentum-strategy performance of small funds, we nonetheless continue to reject the hypothesis that these funds perform as well as on-paper stock portfolios. We can also reject non-monotonicity of momentum compensation across size categories using the bootstrap test of Patton and Timmermann (2010): we find momentum strategy performance is significantly decreasing in fund size. This feature makes intuitive sense in that momentum is a high-turnover strategy, and larger funds suffer greater market impact costs in implementing momentum than smaller funds.

From this analysis we conclude that heterogeneity among asset managers is important when considering the net-of-costs returns to momentum. When we focus only on small mutual funds, net-of-costs compensation to momentum looks quite different from that of the average fund. Which momentum premium is of greater interest hinges on whether the researcher evaluates a broad set

33 Groups are assigned separately for each date with cutoffs in terms of December 2016 dollars. The micro-fund group has $TNA_t < \$10M$ and contains 5.2% of the data. The small-fund group has $\$10M < TNA_t < \$50M$ and contains 22.8% of the data. The medium-fund group has $\$50M < TNA_t < \$250M$ and contains 31.8% of the data. The large-fund group has $\$250M < TNA_t < \$1B$ and contains 22.5% of the data. The mega-fund group has $TNA_t > \$1B$ and contains 17.7% of the data.

Table 10
Fama-MacBeth slopes for stocks and mutual funds, size splits.
This table reports Fama-MacBeth estimates of the compensation for factor exposure for domestic equity mutual funds. Coefficients are the average cross-sectional slopes $\bar{\lambda}_k^g$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$ for each group of mutual funds g,

$$r_{it} = \sum_k \lambda_{kt}^{MF,g} \hat{\beta}_{ik} + \epsilon_{it}, \ t = 1, \ldots, T, \ g = 1, \ldots, 5,$$

where k indexes the four Carhart (1997) factors. We partition mutual funds into four groups based on one-month lagged total net assets (TNA), with TNA cutoffs specified in December 2016 USD. The micro-fund group has $TNA_t < \$10M$, the small-fund group has $\$10M < TNA_t < \$50M$, the medium-fund group has $\$50M < TNA_t < \$250M$, the large-fund group has $\$250M < TNA_t < \$1B$, and the mega-fund group has $TNA_t > \$1B$. First-stage regression estimates include these factors only (columns one to four) and the first principal component of market and funding liquidity proxies (columns five to eight). Liquidity proxies and stock portfolio sets are described in Section 3. λ_{small}^{Δ} is the difference between compensation for factor exposure between the 269 stock portfolios and the small-fund group. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks. The final three rows report the p values of F tests of coefficients being jointly different from zero, of F tests of equality of coefficients, and the Patton and Timmermann (2010) test of non-monotonicity of coefficients.

	Baseline specification				Including liquidity PCs			
	MKT	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
λ_{mega}^{MF}	6.66***	3.11*	1.89	−2.53	6.67***	3.15*	1.94	−2.77
	(2.74)	(1.67)	(1.05)	(−0.75)	(2.75)	(1.66)	(1.08)	(−0.84)
λ_{large}^{MF}	6.85***	2.78	0.90	0.86	6.86***	2.83	0.91	0.04
	(2.78)	(1.54)	(0.52)	(0.31)	(2.79)	(1.54)	(0.53)	(0.02)
λ_{medium}^{MF}	7.02***	2.45	0.90	2.36	7.00***	2.45	0.96	1.76
	(2.87)	(1.41)	(0.52)	(0.92)	(2.86)	(1.37)	(0.55)	(0.68)
λ_{small}^{MF}	7.36***	2.94	1.20	3.40	7.30***	2.55	1.37	2.62
	(2.98)	(1.64)	(0.72)	(1.25)	(2.96)	(1.37)	(0.82)	(0.97)
λ_{micro}^{MF}	7.18***	2.60	−2.68	−0.24	7.18***	2.54	−3.29	−0.04
	(2.94)	(1.11)	(−1.32)	(−0.06)	(2.92)	(1.11)	(−1.59)	(−0.01)
λ_{small}^{Δ}	−0.59	2.26**	−0.26	5.45***	−0.53	2.92**	−0.48	6.22***
	(−1.59)	(2.22)	(−0.34)	(3.32)	(−1.4)	(2.58)	(−0.62)	(3.76)
λ^{MF}	6.98***	2.62	1.01	1.54	6.99***	2.64	0.90	1.28
	(2.86)	(1.51)	(0.59)	(0.63)	(2.87)	(1.51)	(0.53)	(0.52)
$\lambda_i = 0$	0.01***	0.46	0.56	0.11	0.01**	0.52	0.52	0.14
$\lambda_i = \lambda^*$	0.13	0.81	0.46	0.13	0.20	0.83	0.44	0.13
$\Delta\lambda \not\lessgtr 0$	0.01***	0.28	0.20	0.01***	0.01***	0.06*	0.28	0.01***

of firms, as in benchmarking applications, or marginal investors, as in discussions of market efficiency. Intriguingly we find that the largest mutual funds earn the most negative compensation for momentum exposure, suggesting that the firm examined in Frazzini et al. (2015) is exceptional, or that non-mutual fund asset managers have different compensation schedules for factor exposure. More generally, our results reveal that the long-standing disagreement on the profitability of momentum strategies likely arises in part because market-wide and single-firm analyses, e.g., Lesmond et al. (2004) and Frazzini et al. (2015), respectively, focus on different sets of institutions with different factor implementation technologies.

6.2. Implementation costs over time

The preceding analysis considers how implementation costs vary in the cross section. In this section, we investigate determinants of time series variation in implementation costs. Fig. 2 plots the log return of the on-paper factor–mimicking portfolios minus the log return of the corresponding mutual fund factor–mimicking portfolio. To do this we invoke the interpretation of Fama-MacBeth coefficients λ_{kt} as the date t return on a portfolio with a unit loading on factor k and zero loading on all other factors. Our series is the centered rolling difference in performance,

$$y_k(t) = \sum_{s=t-6}^{t+6} \log\left(1+\lambda_{kt}^{S}\right) - \log\left(1+\lambda_{kt}^{MF}\right) \approx \sum_{s=t-6}^{t+6} \lambda_{kt}^{\Delta}. \quad (9)$$

The four panels of Fig. 2 depict factor implementation costs for each set of liquidity proxies using the 269 stock portfolios as the on-paper return benchmark. Although magnitudes vary slightly across specifications, the two slope series are highly similar for each factor. The implementation gap is clearly rank-ordered as *UMD, HML, SMB*, and *MKT*, with large and positive implementation gaps for *UMD* and *HML*, no implementation gap for *SMB*, and a small negative implementation gap for *MKT*. The difference series are also affected by macroeconomic events. All four implementation gaps fall before the end of the tech bubble of the late 1990s and rise during the subsequent crash and the Great Recession of 2007–2009. One interpretation of this feature is that factor returns are most accessible by investment managers when market liquidity is abundant and funding constraints are unlikely to be binding.

Perhaps the most intriguing feature of Fig. 2 is the absence of a clear trend in strategy implementation costs. This feature contrasts with well-documented secular

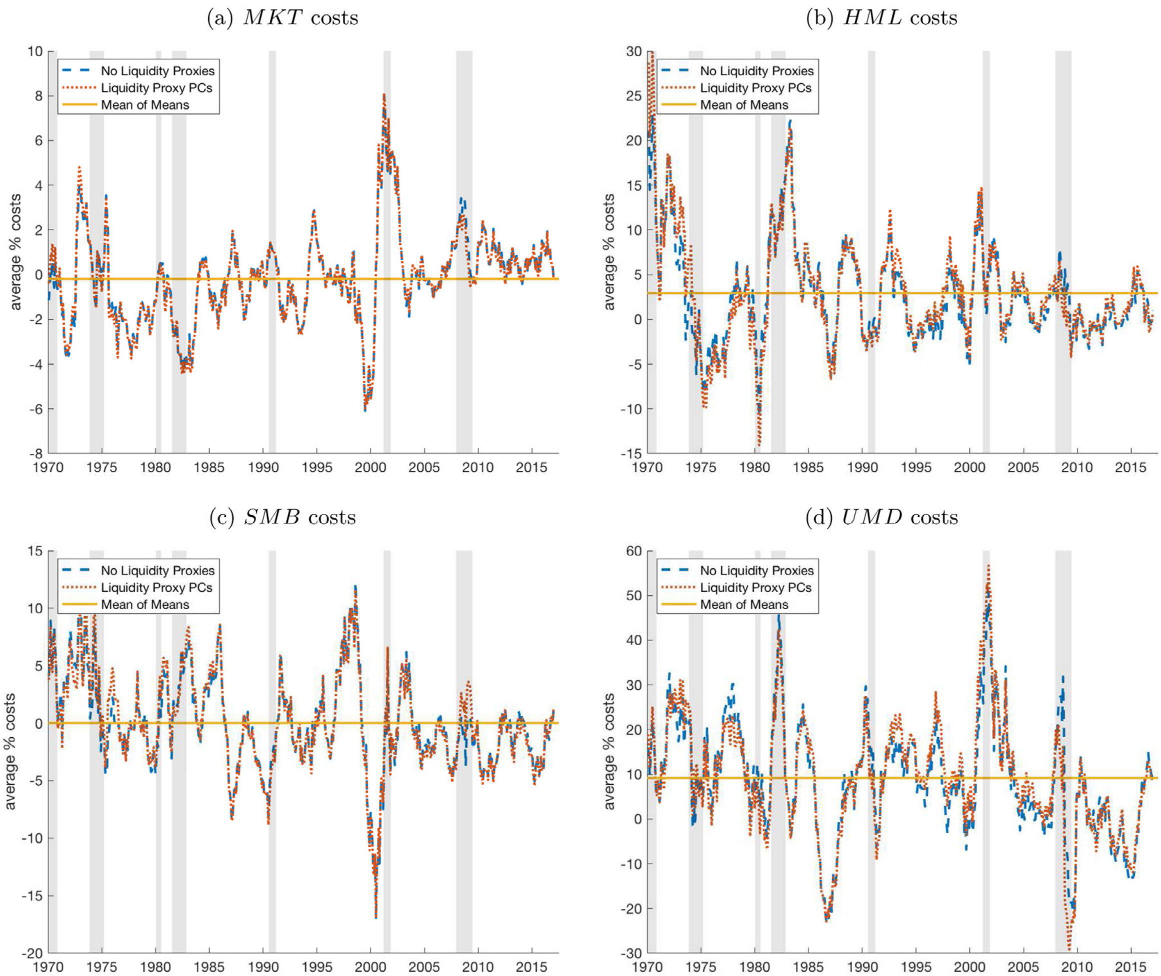

Fig. 2. Rolling performance difference between mutual funds and stocks ($\beta = 1$).
This figure plots the rolling difference between log Fama-MacBeth cross-sectional slopes for stock portfolios (S) and mutual funds (MF). Each series $y_k(t)$ equals the centered rolling difference

$$y_k(t) = \sum_{s=t-6}^{t+6} \log\left(1 + \lambda_{kt}^{S}\right) - \log\left(1 + \lambda_{kt}^{MF}\right),$$

where λ_{kt} are cross-sectional slopes from monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$. Each panel plots differences in slopes when no liquidity proxies are included in the time series regressions and when the first principal components (PCs) of market liquidity proxies and of funding liquidity proxies are included. Stock portfolio slopes are estimated using all 269 portfolios described in Section 3. Solid lines depict averages of series means. NBER recessions are in gray.

declines in bid-ask spreads and commissions since 1970 (e.g., Jones, 2002; Corwin and Schultz, 2012). An equilibrium perspective on the size of the asset management sector suggests why we instead obtain a stationary time series.[34] As trading technology improves and equity intermediation becomes more competitive, the cost of trading the first dollar of a factor strategy declines. Perceived sector-level alphas increase for factor investors, and aggregate inflows attract new entrants (as in Fig. 1) or contribute to the growth of existing fund managers (as in Berk and Green, 2004). These inflows increase the scale of factor investing, which in turn increases non-proportional transactions costs such as price impact. In equilibrium, this process continues until factor alphas fall to zero for the marginal dollar. Consequently, the average dollar invested in factor strategies could see no reduction in implementation costs despite improvements in trading technology.

Following the suggestion of a referee, we also evaluate implementation costs around known changes in trading costs. We consider three key events: the 1975 “May Day” move to negotiated trading commissions, the June 1997 reduction in the minimum tick size to 16ths of a dollar, and the April 2001 decimalization of US stock markets. In selecting months, we choose the date of completion for transitions by the New York Stock Exchange, the dominant exchange of these eras. We use these event dates to partition our sample into four parts, and we estimate average

[34] Augmented Dickey-Fuller tests reject the null of a unit root in implementation costs at the 0.1% significance level in all series.

Table 11
Implementation cost estimates before and after important market structure changes.
This table reports Fama-MacBeth estimates of mutual fund implementation costs before and after three important market structure changes. Implementation costs are measured as the difference in average compensation for factor exposure between mutual fund and stock portfolios. Costs are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$,

$$r_{it} = \sum_k \lambda^S_{kt} \hat{\beta}_{ik} 1_{i \in S} + \sum_k \lambda^{MF}_{kt} \hat{\beta}_{ik} 1_{i \in MF} + \epsilon_{it}, \ t = 1, \ldots, T,$$

where k indexes the four Carhart (1997) factors and λ^Δ is defined as $\lambda^S - \lambda^{MF}$. Event dates are the 1975 "May Day" move to negotiated trading commissions, the 1997 reduction in the minimum tick size to 16ths of a dollar, and the 2001 decimalization of US stock markets. Stock portfolio sets are described in Section 3. All coefficients are annualized and reported in percent. t statistics for differences in implementation costs are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

	Baseline specification		Including liquidity PCs	
	$N_S = 100$	$N_S = 269$	$N_S = 100$	$N_S = 269$
Panel A: Value-weighted stock portfolios				
$\bar{\lambda}^\Delta_{1970-1974}$	6.24	6.09	7.13	6.93
$\bar{\lambda}^\Delta_{1975-1996}$	2.91	2.38	2.72	2.22
$\bar{\lambda}^\Delta_{1997-2000}$	2.37	1.91	3.65	3.05
$\bar{\lambda}^\Delta_{2001-2016}$	1.34	1.29	1.44	1.37
$\bar{\lambda}^\Delta_{1975-1996} - \bar{\lambda}^\Delta_{1970-1974}$	−3.34**	−3.71***	−4.41***	−4.71***
	(−2.31)	(−2.74)	(−2.93)	(−3.32)
$\bar{\lambda}^\Delta_{1997-2000} - \bar{\lambda}^\Delta_{1975-1996}$	−0.54	−0.47	0.93	0.83
	(−0.32)	(−0.3)	(0.54)	(0.52)
$\bar{\lambda}^\Delta_{2001-2016} - \bar{\lambda}^\Delta_{1997-2000}$	−1.03	−0.62	−2.21	−1.68
	(−0.63)	(−0.41)	(−1.39)	(−1.12)
$F\left(\bar{\lambda}^\Delta_{1970-1974} = \ldots = \bar{\lambda}^\Delta_{2001-2016}\right)$	3.73*	4.08**	4.92**	5.37**
p-value	0.05	0.04	0.03	0.02
Panel B: Equal-weighted stock portfolios				
$\bar{\lambda}^\Delta_{1970-1974}$	6.73	6.89	7.61	7.90
$\bar{\lambda}^\Delta_{1975-1996}$	3.50	3.55	3.19	3.31
$\bar{\lambda}^\Delta_{1997-2000}$	3.22	3.23	4.39	4.36
$\bar{\lambda}^\Delta_{2001-2016}$	1.94	2.57	2.00	2.71
$\bar{\lambda}^\Delta_{1975-1996} - \bar{\lambda}^\Delta_{1970-1974}$	−3.23**	−3.33**	−4.42***	−4.59***
	(−2.15)	(−2.30)	(−2.85)	(−3.08)
$\bar{\lambda}^\Delta_{1997-2000} - \bar{\lambda}^\Delta_{1975-1996}$	−0.27	−0.33	1.20	1.05
	(−0.16)	(−0.19)	(0.68)	(0.60)
$\bar{\lambda}^\Delta_{2001-2016} - \bar{\lambda}^\Delta_{1997-2000}$	−1.28	−0.66	−2.39	−1.65
	(−0.79)	(−0.39)	(−1.51)	(−0.95)
$F\left(\bar{\lambda}^\Delta_{1970-1974} = \ldots = \bar{\lambda}^\Delta_{2001-2016}\right)$	3.41*	2.78*	4.68**	4.03**
p-value	0.07	0.10	0.03	0.05

implementation costs across the four factors in each. This composite represents the typical all-in cost faced by mutual funds per unit of risk exposure. We then test for whether the differences in costs are statistically significant for each event and jointly across events.

Table 11 reports cost estimates and statistical tests for whether their differences are different from zero before and after market structure changes. The table reveals a clear decrease in implementation costs across market structure changes, and costs are monotone decreasing or nearly so (the point estimate $\bar{\lambda}_{1997-2000}$ is sometimes larger than $\bar{\lambda}_{1975-1996}$). The corresponding point estimates are typically negative in the second subpanel, but these differences are statistically different from zero only for the May Day shift to negotiated commissions. Typically, the implementation cost series is too short and too noisy to detect differences reliably, as we observe in Fig. 2. We reject the hypothesis that implementation costs are equal across market structures at the 10% level without liquidity proxies and at the 5% level with them. We conclude that implementation costs indeed decline with important changes in trading costs. That these differences are not even larger suggests that trading-cost reductions are counteracted by another equilibrating mechanism, such as the previously conjectured growth in fund size generally or dollar-weighted factor exposures specifically.

Our conjectured equilibrium adjustment mechanism hinges on non-proportional trading costs, and it generates a testable prediction that industry-level inflows increase implementation costs of factor strategies. We analyze this relation between implementation costs, flows, and illiquidity more formally by relating the cost time series with liquidity and fund flow proxies. We start by constructing illiquidity proxies as the first principal components of market

Table 12
Liquidity, flows, and the implementation gap.
This table reports estimates of a regression of annualized implementation gaps λ^{Δ}_{kt} on the first principal component (PC) of funding liquidity proxies (*FL*) and market liquidity proxies (*ML*), as well as the cross-sectional average and standard deviation of fund inflows (*MFLOW* and *DFLOW*, respectively),

$$\lambda^{\Delta}_{kt} = \alpha + \beta_{MFLOW} MFLOW + \beta_{DFLOW} DFLOW + \beta_{ML} PC_{ML} + \beta_{FL} PC_{FL} + \epsilon_{kt},$$

where k indexes the four Carhart (1997) factors. Implementation costs are estimated with principal components of market and funding liquidity proxies in the time series regressions for mutual fund and stock portfolio betas. Liquidity proxies and stock portfolio sets are described in Section 3. Illiquidity principal components have unit standard deviation and are constructed to be negatively correlated with the market return and positively correlated with the VIX/VXO. Fund flows are $TNA_{it}/TNA_{i,t-1} - (1 + R_{it})$. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

	MKT			HML			SMB			UMD		
Panel A: Baseline specification												
β_{MFLOW}	0.07		−0.02	1.91**		1.30	−0.02		−0.19	0.65		−0.03
	(0.27)		(−0.06)	(2.21)		(1.47)	(−0.04)		(−0.34)	(0.50)		(−0.03)
β_{DFLOW}	−0.05		−0.10	−2.16*		−2.45**	0.13		0.09	−1.81		−2.09
	(−0.13)		(−0.27)	(−1.87)		(−2.15)	(0.23)		(0.17)	(−0.92)		(−1.08)
β_{ML}		0.53*	0.55*		2.80***	3.03***		0.44	0.44		2.75	3.03
		(1.65)	(1.68)		(2.81)	(3.31)		(0.59)	(0.59)		(1.53)	(1.64)
β_{FL}		−0.46	−0.49		−3.08***	−3.08***		−0.68	−0.73		−2.92	−0.73
		(−1.34)	(−1.35)		(−3.28)	(−3.12)		(−0.88)	(−0.88)		(−1.57)	(−0.88)
α	−0.21	−0.21	−0.21	2.59***	2.59***	2.59***	−0.07	−0.07	−0.07	7.30***	7.30***	7.30***
	(−0.88)	(−0.88)	(−0.88)	(3.85)	(3.87)	(3.92)	(−0.14)	(−0.14)	(−0.14)	(5.59)	(5.57)	(5.65)
Panel B: Including liquidity principal components												
β_{MFLOW}	0.10		−0.02	2.38**		1.55	−0.14		−0.25	1.87		0.00
	(0.39)		(−0.09)	(2.50)		(1.64)	(−0.29)		(−0.46)	(1.42)		(0.00)
β_{DFLOW}	−0.02		−0.07	−1.95		−2.28*	0.01		−0.07	−1.50		−1.70
	(−0.05)		(−0.19)	(−1.57)		(−1.90)	(0.02)		(−0.12)	(−0.76)		(−0.91)
β_{ML}		0.57*	0.58*		3.42***	3.61***		0.76	0.79		2.38	2.61
		(1.82)	(1.83)		(3.38)	(3.88)		(0.90)	(0.93)		(1.41)	(1.51)
β_{FL}		−0.61*	−0.63*		−4.12***	−4.01***		−0.54	−0.64		−6.69***	−0.64
		(−1.79)	(−1.77)		(−4.28)	(−4.05)		(−0.69)	(−0.76)		(−3.77)	(−0.76)
α	−0.22	−0.22	−0.22	2.83***	2.83***	2.83***	−0.02	−0.02	−0.02	7.55***	7.55***	7.55***
	(−0.91)	(−0.92)	(−0.92)	(3.92)	(4.03)	(4.08)	(−0.03)	(−0.03)	(−0.03)	(5.73)	(5.92)	(5.98)

liquidity proxies and of funding liquidity proxies, as described in Section 4.3. We also construct flow variables to capture costs associated with fund inflows and outflows. Fund flows are the component of asset growth not explained by returns,

$$flow_{it} = \frac{TNA_{it}}{TNA_{i,t-1}} - (1 + R_{it}). \tag{10}$$

We summarize the distribution of flows with its first and second cross-sectional moments, the cross-sectional average fund flow (*MFLOW*) and the cross-sectional dispersion in fund flows (*DFLOW*). In addition to speaking to returns to scale, flow variables are a natural candidate for explaining trading costs because large flows into the mutual fund sector or reshuffling of assets among mutual funds generates liquidity demands. To enhance interpretability, we normalize all right-hand-side variables to have mean zero and standard deviation one.

Table 12 reports results from regressions of λ^{Δ}_{kt}, for each of the factors $k = 1, 2, 3, 4$, on the liquidity and fund flow proxies,

$$\lambda^{\Delta}_{kt} = \alpha + \beta_{MFLOW} MFLOW + \beta_{DFLOW} DFLOW + \beta_{ML} PC_{ML} + \beta_{FL} PC_{FL} + \epsilon_{kt}. \tag{11}$$

We report only value-weighted results for the 269 stock portfolios because relations between costs and liquidity proxies are similar for value-weighted and equal-weighted stock portfolios and for 100 and 269 stock portfolios. Columns (1)-(3) of Table 12 refer to the market factor, columns (4)-(6) refer to the value factor, columns (7)-(9) refer to the size factor, and columns (10)-(12) refer to the momentum factor.

We draw four lessons from Table 12. First, the time-invariant component of implementation costs from Eq. (4) is large and positive for these factors, as evidenced by the constant terms for *HML* and *UMD*. Second, focusing on flows, average inflows are weakly associated with higher implementation costs for value and momentum factors, and cross-sectional dispersion in flows is weakly associated with lower implementation costs for these factors. We find no flow-cost relations for market and size factors, as is expected because these costs are small in magnitude to begin with. We interpret these relations as suggestive evidence that inflows are expensive from a transactions-cost standpoint for funds trading value and momentum strategies, thereby contributing to diseconomies of scale and stationary average implementation costs, and that reallocation of funds within the mutual fund sector may increase liquidity trading [in a Kyle (1985) sense], thereby reducing average transactions costs for value and momentum traders. Third, focusing on illiquidity principal components, market illiquidity increases implementation costs, and particularly so for value and momentum strategies. Intuitively trading becomes more expensive when market liquidity is low. Fourth, funding illiquidity decreases implementation costs (again most strongly for *HML* and *UMD*). We conjecture that mutual funds are insulated from funding liquidity shocks that affect highly levered institutional asset managers such as hedge funds [Sadka (2010) and Boyson et al. (2010), among others, discuss hedge funds' particular vulnerability to funding liquidity shocks], and, hence, mutual funds can acquire the ingredients of factor strategies from distressed asset managers at a discount during times of strained funding liquidity.

7. Conclusion

Existing methods for assessing the implementation costs of financial market anomalies use proprietary trading data for single firms or market-wide trading data combined with parametric transactions cost models. We propose an extension of Fama-MacBeth regression to estimate implementation costs using only returns data from stocks and mutual funds. Doing so frees us from the requirement of specifying factor trading strategies and transaction costs models that may be incomplete or misspecified. Moreover, the ready availability of returns data for a large universe of investment managers enables detailed investigation of factor implementation costs in the cross-section and over time.

We demonstrate that mutual funds are generally poorly compensated for exposure to some common risk factors. Our estimates based on Fama-MacBeth regressions imply that implementation costs erode almost the entirety of the return to value and momentum strategies for typical mutual funds but have little effect on market and size factor strategies. Taken together, these results paint a sobering picture of the real-world returns to the most important financial market anomalies. These costs derive in part from institutional constraints often ignored in studies of academic anomalies, such as shorting and investability constraints, but even these frictions do not fully explain high observed costs. We find suggestive evidence that unprofitable deviations from standard academic strategies and greater market impact associated with fund growth contribute to the remainder.

Sample splits reveal considerable heterogeneity in implementation costs among funds. Using a very different estimation method and data set, our results agree with the Lesmond et al. (2004) analysis that momentum profits in particular may be out of reach for a typical asset manager. However, we find smaller funds and funds that better track academic factors tend to perform significantly better in earning factor premia than larger funds and funds with greater tracking error. In this respect, markets may be efficient from the perspective of an average mutual fund, even if some segments of the mutual fund space see a very different picture of risk and return net-of-costs. Analyses using proprietary data from single funds cannot reveal such heterogeneity.

The nonparametric, market-based method for estimating all-in implementation costs proposed in this paper can be viewed as an independent check on prior work because it differs in both estimation strategy and data employed. The assumptions underlying our approach are few and transparent, and our stark findings on realizable factor premia obtain under a wide range of alternative specifications. While we do not anticipate resolving a decades-old dispute on whether momentum is accessible to typical investors, our approach forces a conversation about the palatability of assumptions on representativeness, price impact, and the like made in existing work.

Appendix A. Mutual fund filters

We clean the CRSP mutual fund database at the individual fund and fund group levels. We first clean at the lowest level of aggregation to deal with missing and erroneous data, and then we filter our sample based on fund group-level information.

A.1. Cleaning procedures at the fund level

We first flag fund-dates with reporting frequencies less than monthly in the monthly returns file (6,526,393 observations). As discussed by Elton et al. (2001) and Fama and French (2010), about 15% of funds before 1983 report returns only annually, and we mark as missing the fund returns for which neither adjacent month has a non-zero and non-missing return. These annual returns comprise 1.71% of fund-month observations.

Next we construct current and lagged total net asset values for value-weighting fund returns within and across fund groups. Nearly a tenth of TNA values are undefined, and we interpolate TNA values to avoid discarding such a large fraction of the data. Before interpolating, we flag as missing invalid TNA values that arise because of recording errors or bottom coding. As noted in the CRSP mutual fund database documentation, entries of $100,000 denote TNAs of less than or equal to this value. Although not documented, entries of $1000 seem to serve a similar role. We eliminate bottom-coded TNAs by setting to missing values less than or equal to $100,000 US dollars. Likewise, we set to missing TNA values exceeding $1 trillion US dollars, as no single fund has ever reached this value. Imposing these filters, 14.9% of TNA observations are flagged as missing.

We interpolate TNAs in three steps. First, we compute a predicted TNA by multiplying the last available TNA value by cumulative returns since that date. This predicted TNA value misses inflows and outflows from the fund. Second, when available, we reconcile predicted TNAs and the next filled TNA observation. The ratio of true TNA to predicted TNA (minus one) is a discrepancy associated with fund inflows or outflows. We assume flows are constant between known TNA values, and we multiply predicted TNA by $(1 + discrepancy)^{s/\Delta t}$, where s is the number of months since the last known TNA value and Δt is the number of months between TNA values. We assume a discrepancy of zero if there is no next known TNA. Third, we run the first and second steps backward to use return data to fill in TNAs before the first reported TNA value. Given the interpolated values, we again set as missing any TNA values smaller than $100,000 or greater than $1 trillion. The filling and cleaning procedures reduce the number of missing TNA values to 2.8% of the data.

Share classes differ from one another in their fee structures, and we account for this variation before aggregating across share classes within a fund. We convert net re-

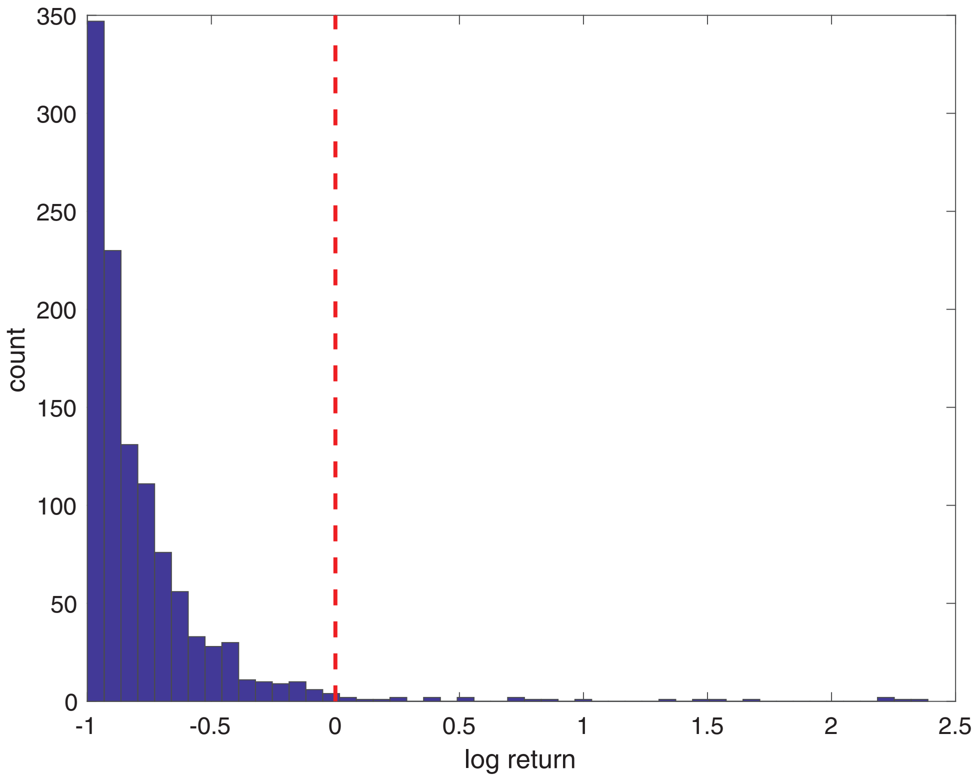

Fig. A.1. Distribution of fund returns.
This figure plots the distribution of the log of absolute monthly mutual fund returns. We truncate the plot to −1 on the left to maintain resolution on the extreme returns on the right. The dashed line represents of cutoff at $|r| = 100\%$.

turns to gross returns by adding to net returns the annual expense ratio divided by 12, following Fama and French (2010). The fund summary file has missing or non-positive expense ratios for 16.9% of observations, and we take several steps to fill in the missing data. First, as before, we fill missing expense ratios with the nearest observation with a non-missing value within each CRSP fund number group. This operation reduces the number of missing expense ratios to 8.4% of the summary data. We then merge the monthly return data with the summary data by fund number and calendar quarter. This merge assigns expense ratios to 76.2% of fund-month observations. For unmerged observations, we merge again on fund number and year, where we take the average expense ratio within the fund number-year in the fund summary file. This operation boosts the number of fund-month observations with an expense ratio to 88.5% of the data or 5,774,820 observations. We then drop the 89 observations with expense ratios exceeding 50% as these are almost certainly data errors.

We next filter out extreme return observations resulting from data errors. For example, we do not wish to include the recorded return of 533% on the Deutsche Equity 500 Index Fund in September 1997. Berk and van Binsbergen (2015) and Pastor et al. (2015) address these errors in part using external Bloomberg and Morningstar databases. We take a simpler approach to eliminate errors. We drop the 23 observations with reported returns exceeding one (i.e., 100%) in absolute value. This approach is inspired by the shape of the tail of extreme returns in the data depicted in Fig. A.1: the frequency of extreme returns decays roughly exponentially until $|r| = 100\%$, with a smattering of randomly spaced returns beyond this value. These observations appear to come from a different distribution, and for this reason, we classify them as likely errors.

Because our analysis concerns mutual funds, we filter out exchange-traded funds (ETFs), exchange-traded notes (ETNs), and variable annuity underlying (VAU) funds. To do this, we discard any observations for which "et_flag" indicates an ETF or ETN or "vau_fund" indicates a VAU at any time in a fund's life. These exclusions total 9.1% of observations.

A.2. Aggregation into fund groups

Having accounted for the salient variation across share classes, we next identify share classes of the same fund. As a preliminary step, we fill missing fund names using the nearest observation with a non-missing fund name within each CRSP fund number group. Of the 1,859,702 observations in the fund summary file, we assign fund names for 19,460 observations and remove 242 observations without recoverable names. We then repeat this procedure for missing fund tickers. This matching step assigns 116,238 of the 274,875 observations with missing tickers. By contrast with observations missing names, we retain observations with missing tickers.

We then follow almost exactly the two-step fund class grouping procedure of Berk and van Binsbergen (2015) and Pastor et al. (2015).

1. We identify share classes following three mutually exclusive rules. First, if the CRSP fund name contains a semicolon and the phrase after the last semicolon does not contain a forward slash, we retain the fund name prior to the last semicolon as the fund group name. Second, if the CRSP fund name contains a forward slash, and the entire phrase after the last forward slash does not contain a space or a semicolon, we use the word prior to the last forward slash as the fund group name. Third, if neither rule applies, we assume that the CRSP fund name does not include a share class.

   We make a minor adjustment to their methodology before applying these rules. Although handling semicolons is straightforward, forward slashes—the other class-name delimiter used in CRSP—require more care. For example, fund names include “Franklin/Templeton” and “M/M” (money market), so “/” does not serve only as a delimiter, and the absence of a space does not guarantee that the subsequent string is a share class. So as a preliminary step, we replace forward slashes in T/F, T/E, M/M, L/S, Small/Mid, Long/Short, S/T, and L/T with backslashes in fund names.
2. We define equivalent funds as those sharing an adjusted name or a ticker symbol. To do so, we iteratively build equivalence classes of funds with equivalent names and/or ticker symbols. Because equivalence is transitive, a pair of funds that shares a name, and another fund that shares a ticker with the second fund, are all considered to be of the same group.

This mapping reduces the 61,734 surviving unique fund identifications (IDs) in the CRSP monthly returns file to 23,613 unique fund groups. Of the 6,522,095 observations in the monthly return file, only 4298 of these are not assigned a fund group, and these observations are dropped.

A.3. Cleaning procedures at the fund group level

We construct fund group returns and total net assets by taking a weighted average of returns across component fund IDs. The return weights are one-month lagged TNAs. We retain observations for which the lagged TNA is undefined but the fund group only has one fund ID, that is, the one fund ID has an effective weight of 100%. Likewise, fund group TNAs are the sum of current TNA values across component fund IDs. Aggregating funds across share classes delivers 2,244,101 monthly fund-group observations.

As Fama and French (2010) note, “incubation bias arises because funds typically open to the public—and their prerelease returns are included in mutual fund databases—only if the returns turn out to be attractive.” We follow their approach to countering incubation bias by keeping observations only after a fund group achieves a TNA of at least $10 million (in December 2016 dollars).[35] We retain data from funds that later drop below this threshold to avoid introducing a selection bias. Dropping fund groups that never achieve a $10 million TNA eliminates 2.3% of fund group-month observations. Dropping observations from potential incubation periods before the $10 million threshold is achieved eliminates another 3.9% of the sample.

Next, we filter fund groups based on fund name and objective. We first exclude all funds with names containing “ETF,” “ETN,” “exchange-traded fund,” “exchange traded fund,” “exchange-traded note,” “exchange traded note,” “iShares,” and “PowerShares” (not case sensitive) as a redundant filter on top of the CRSP-based ETF/ETN filter. These exclusions eliminate 3006 observations. We then exclude any funds with names that have clear international or non-equity connotations: “international,” “intl,” “bond,” “emerging,” “frontier,” “rate,” “fixed income,” “commodity,” “oil,” “gold,” “metal,” “world,” “global,” “China,” “Europe,” “Japan,” “real estate,” “absolute return,” “government,” “exchange,” “euro,” “India,” “Israel,” “treasury,” “Australia,” “Asia,” “pacific,” “money,” “cash,” “yield,” “U.K.,” “UK,” “kingdom,” “municipal,” “Ireland,” “LIBOR,” “govt,” “obligation,” “money,” “cash,” “yield,” “mm,” “m/m,” “diversified” (but not “diversified equity”), and “short term” (not case sensitive). This filter complements our requirement that a fund have a domestic equity “ED” CRSP objective code.[36] These filters reduce the number of valid funds from 12,691 to 4282, and the corresponding number of non-missing return observations decreases to 740,899 for the entire December 1961 to December 2016 CRSP mutual fund database.

Finally, we restrict the set of funds to those with at least two years of monthly data in our 1970–2016 sample period. This filter reduces our sample to 4267 mutual funds with 724,995 non-missing return observations. Summary statistics for this sample are reported in Table 1.

Appendix B. Quality of factor approximation by mimicking portfolios

Fama-MacBeth cross-sectional slopes are factor-mimicking portfolio returns. This equivalence allows us to interpret differences in slopes as differences in factor compensation for stock and mutual fund portfolios. In standard Fama-MacBeth regressions in which a constant is included in the cross-sectional step, mimicking portfolio weights and loadings on other factors are constrained to zero in synthesizing the mimicking return on factor k. In our setting, the cross-sectional slopes have the same interpretation of mimicking portfolio returns, but the zero-weight constraint is not enforced. Moreover, because our mutual fund panel is unbalanced, the implied weights of

[35] Our inflation index is the Consumer Price Index for All Urban Consumers (CPIAUCSL) series provided by the Federal Reserve Bank of St. Louis' FRED database.

[36] The CRSP objective code unifies Wiesenberger objective codes for 1962–1993 data, Strategic Insight objective codes for 1993–1998 data, and Lipper objective codes for 1998–2016 data.

Table A.1
Time series regressions of factor-mimicking portfolio returns on Carhart factors.
This table reports coefficients and R^2s of time series regressions of factor-mimicking portfolio returns on the four Carhart (1997) factors, that is,

$$\hat{\lambda}_{kt} = \gamma_0 + \gamma_1 MKT_t + \gamma_2 HML_t + \gamma_3 SMB_t + \gamma_4 UMD_t + \epsilon_{kt},$$

for $k \in \{MKT, HML, SMB, UMD\}$. Panels A and B use mimicking returns from our baseline specification [Eqs. (1) and (2)], and Panels C and D use mimicking returns from the liquidity principal-component augmented specification [Eqs. (6) and (2)]. We suppress reporting of the constant term γ_0 because it is identically zero in all specifications. "Stock"–mimicking portfolios use the expansive portfolio set of 269 stock portfolios, and "MF" mimicking portfolios use all domestic equity mutual funds. Liquidity proxies and stock portfolio sets are described in Section 3.

	MKT		*HML*		*SMB*		*UMD*	
	Stocks	MFs	Stocks	MFs	Stocks	MFs	Stocks	MFs
Panel A: Baseline specification, 1970–2016								
MKT	**1.00**	**1.00**	0.00	0.04	0.00	−0.01	0.00	0.06
HML	0.00	−0.02	**1.00**	**0.91**	0.00	0.00	0.00	−0.05
SMB	0.00	0.03	0.00	0.04	**1.00**	**0.99**	0.00	−0.03
UMD	0.00	0.01	0.00	−0.01	0.00	0.00	**1.00**	**0.95**
R^2	0.99	0.99	0.79	0.71	0.94	0.89	0.95	0.74
Panel B: Baseline specification, 1993–2016								
MKT	**1.00**	**1.00**	0.00	0.02	0.00	−0.01	0.00	−0.01
HML	0.00	0.00	**1.00**	**0.97**	0.00	0.00	0.00	−0.03
SMB	0.00	0.01	0.00	0.00	**1.00**	**1.01**	0.00	0.00
UMD	0.00	0.00	0.00	−0.01	0.00	0.00	**1.00**	**0.95**
R^2	0.99	0.99	0.74	0.75	0.95	0.93	0.96	0.79
Panel C: Including liquidity principal components, 1970–2016								
MKT	**1.00**	**0.99**	0.01	0.04	−0.01	−0.01	0.00	0.02
HML	0.00	−0.02	**1.02**	**0.93**	0.00	−0.01	0.01	−0.08
SMB	0.00	0.03	0.00	0.04	**1.00**	**0.99**	0.00	−0.02
UMD	0.00	0.01	0.01	0.01	0.00	0.00	**1.00**	**0.94**
R^2	0.99	0.99	0.79	0.70	0.94	0.90	0.95	0.74
Panel D: Including liquidity principal components, 1993–2016								
MKT	**1.00**	**0.99**	0.00	0.00	0.01	0.00	−0.01	−0.06
HML	0.00	0.00	**1.01**	**0.99**	0.00	−0.01	0.00	−0.08
SMB	0.00	0.00	0.00	0.00	**1.00**	**1.01**	0.00	0.00
UMD	0.00	0.00	0.01	0.01	0.00	0.00	**1.00**	**0.92**
R^2	0.99	0.99	0.74	0.75	0.95	0.93	0.96	0.80

any given mutual fund in the best approximating portfolio are likely to vary over time.

We now confirm that relaxing the zero- and constant-weight constraints does not affect the mimicking portfolio's essential properties of unconditional betas equal to one on the mimicked factor, unconditional betas equal to zero on other factors, and no average compensation not attributable to the factors themselves (e.g., no contamination by risk-free compensation). If these conditions are met, the return on the mimicking portfolios retains the key interpretation as the return to factor k for the selected set of test assets.

We verify these properties using time series regressions of factor-mimicking portfolio returns on the Carhart factor series,

$$\hat{\lambda}_{kt} = \gamma_0 + \gamma_1 MKT_t + \gamma_2 HML_t + \gamma_3 SMB_t + \gamma_4 UMD_t + \epsilon_{kt}, \tag{B.1}$$

for $k \in \{MKT, HML, SMB, UMD\}$. If the mimicked portfolios were poor approximations, for example, they have the wrong scale, the coefficient on factor k would differ from one and/or the coefficients on the other factors or the constant would differ from zero. Note that coefficients may differ slightly from one or zero because our time series is of finite length, and some periods' mimicking returns may be noisy approximations because of imperfect spanning of the academic factors. The variance of the error term is likely to be especially large in earlier eras in which mutual funds may have small loadings on momentum or other academic factors.

Table A.1 reports results from the regression in Eq. (B.1). The constant γ_0 is not reported because it is zero to two decimal places in all specifications. Because the stock portfolios used in constructing the mimicking portfolios contain those used to construct the academic factors, coefficients are exactly one and zero for stock portfolios in the baseline specification. Coefficients for mutual fund portfolios differ slightly from one and zero, as expected, but these differences are economically negligible, and factor k variation dominates in explaining $\hat{\lambda}_{kt}$. The greater noise in the mutual-fund–mimicking portfolios manifests also in the slightly lower R^2s of the mutual-fund regressions relative to the stock portfolio regressions. These differences are small for all factors other than momentum, for which mutual funds historically have smaller dispersion in loadings.

Results are virtually identical in the lower panel with one important difference. Because the first- and second-stage regressions differ in the principal-component augmented regressions, coefficients need not be exactly one and zero in Eq. (B.1), even for the stock portfolios that would otherwise span the factors perfectly. Coefficients typically differ from zero or one by less than 2%, with the largest differences being for mutual-fund momentum at 8%. Still, the scale differences in the mimicking portfolios do not come close to explaining differences in factor means for stock portfolios and mutual funds: even an 8% (1%) scale difference for momentum cannot explain the nearly 100% (50%) difference in momentum (value) premia between stock portfolios and mutual funds.

Appendix C. Bias of symmetric Fama-MacBeth regressions with general h_{it}

Section 4.3 implements an asymmetric Fama-MacBeth regression in which the first stage includes liquidity proxies, and the second stage does not. If instead we were to also include the loadings on the liquidity proxies in Eq. (2), the second-stage regression becomes

$$r_{it} = \sum_k \lambda^S_{kt}\hat{\beta}_{ik}1_{i\in S} + \sum_k \lambda^{MF}_{kt}\hat{\beta}_{ik}1_{i\in MF} + \sum_l \lambda^S_{lt}\hat{\gamma}_{il}1_{i\in S} + \sum_l \lambda^{MF}_{lt}\hat{\gamma}_{il}1_{i\in MF} + \epsilon_{it}.\ t = 1, \ldots, T. \tag{C.1}$$

From the conjectured return process of Eq. (5), $\hat{\lambda}^S_t = \lambda^S_t$, $\hat{\lambda}^{\Delta}_{kt} = \bar{\eta} + \frac{cov((\eta_i-\bar{\eta})\beta_i,\beta_i)}{var(\beta_i)}$, and $\hat{\lambda}^{\Delta}_{lt} = \eta_{lt}$. The problem with this approach is that the $\hat{\lambda}^{\Delta}_{lt}$ terms absorb the time-varying part of η_{it}, so we can no longer cleanly attribute time-varying costs to each return factor. Moreover, the logic of mutual funds scaling down strategies in the face of high costs applies to η_{it} rather than to η_i.

To resolve the first issue, we need to decompose η_{st} into factor-specific parts. The sum of all time-varying costs is

$$TVC_{it} \equiv \sum_l \eta_{lt}\gamma_{il} = \sum_l \eta_{lt}\left(\sum_k \gamma_{ikl}\beta_{ik}\right). \tag{C.2}$$

Regressing total time-varying costs on β_is decomposes costs into factor-specific time-varying parts,

$$TVC_{it} = \sum_t \sum_k \eta_{kt}\beta_{ik}1_t + \epsilon_{it}. \tag{C.3}$$

This regression can be interpreted as projecting time-varying liquidity costs onto the factor-exposure space. However, this rotation is imperfect because of cross-sectional variation in γ_is. To see why dispersion in γ_i is problematic, consider a single coefficient estimate in a one-return factor case of Eq. (C.3),

$$\hat{\eta}_t = \frac{cov(\sum_l \eta_{lt}\gamma_{il}, \beta_i)}{cov(\beta_i)} = \sum_l \eta_{lt}\bar{\gamma}_l + \sum_s \frac{cov(\beta_i\eta_{lt}(\gamma_{il}-\bar{\gamma}_l), \beta_i)}{cov(\beta_i)}. \tag{C.4}$$

The first term represents the average exposure to liquidity factors multiplied by the factors' time t realizations. This is the term of interest, but instead we identify this term plus a cross-sectional bias term.

Focusing on the bias for each l, we might expect higher-than-average cost-factor sensitivities $\gamma_{il} > \bar{\gamma}_l$ to be associated with lower betas if firms are risk averse. Although we would expect betas to be negatively associated with total costs per unit of risk exposure η_{it}, it is not clear what relation the time-varying component alone should have with betas. Because of this ambiguous sign and the additional complexity of this approach, it is preferable not to include the liquidity exposures in the cross-sectional regression step.

Appendix D. Spanning variation in η using many liquidity proxies

Including more covariates increases the likelihood that we span variation in implementation costs, η_{it}, by including all salient liquidity proxies. At the same time, including additional highly correlated cost proxies may overfit the first-stage regression and deliver nonsensical cross-sectional slopes in the second stage.

Sparse regression techniques offer a solution to this challenge. We supplement the standard first-stage regression with a Lasso or l_1-penalized regression (Tibshirani, 1994). We augment the least-squares minimization problem in the time series regressions with additional terms to penalize liquidity coefficients,

$$\min_{\beta,\tilde{\gamma}} \frac{1}{T_i}\sum_t \left(r_{it} - \sum_k f_{kt}\beta_{ik} - \sum_l \tilde{\eta}_{lt}\tilde{\gamma}_{il}\right)^2 + \kappa\left(\sum_k \omega_k|\beta_{ik}| + \sum_l \omega_l|\tilde{\gamma}_{il}|\right), \tag{D.1}$$

where κ represents a penalty term for coefficients different from zero, and ω_k and ω_l represent additional relative penalties explained below. The problem reduces to least squares when $\kappa = 0$; otherwise, liquidity coefficients are compressed toward zero. Note that we do not require a penalization in the cross-sectional step because the second-stage regression omits liquidity proxies. As before, we normalize all liquidity proxies to give them similar scales and an equal chance of entering the Lasso regression.[37]

Lasso simultaneously prevents overfitting in the time series regressions by shrinking coefficients and selects covariates by zeroing out coefficients that would otherwise be close to zero. Both features facilitate the use of many liquidity proxies even when a mutual fund is relatively short-lived. Moreover, we no longer need to choose which measure(s) best approximate the costs faced by each fund, and different liquidity measures can be more salient for different mutual funds. First-stage penalization also knocks out spurious strategy loadings for funds that take on risk exposures unintentionally; a small non-zero loading taken en route to implementing a different strategy will be zeroed out.

The original Lasso implementation sets $\omega_k = \omega_l = 1$ for all k and l. Unfortunately Lasso is not guaranteed to de-

[37] We interpolate missing elements of the VXO/TED series using their matrix-completed values $\phi'_{VXO}g_{ML}$ and $\phi'_{TED}g_{FL}$ from the PCA-ALS procedure described in footnote 22.

Table A.2
Implementation cost estimates, liquidity Lasso.
This table reports Fama-MacBeth estimates of the compensation for factor exposure for stock portfolios (third and fourth rows), domestic equity mutual funds (fifth row), and their difference (first two rows). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$,

$$r_{it} = \sum_k \lambda_{kt}^S \hat{\beta}_{ik} 1_{i\in S} + \sum_k \lambda_{kt}^{MF} \hat{\beta}_{ik} 1_{i\in MF} + \epsilon_{it},\ \ t = 1, \ldots, T,$$

where k indexes the four Carhart (1997) factors and λ^{Δ} is defined as $\lambda^S - \lambda^{MF}$. First-stage regression estimates include these factors and all market and funding liquidity proxies in an adaptive Lasso regression with portfolio-specific penalty parameters chosen by ten-fold cross validation. Liquidity proxies and stock portfolio sets are described in Section 3. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1970 – 2016				1993 – 2016			
	N_S	MKT	HML	SMB	UMD	MKT	HML	SMB	UMD
Panel A: Value-weighted stock portfolios									
λ^{Δ}	100	−0.22	4.97***	0.09	8.71***	−0.02	3.85***	−0.36	6.67***
		(−0.71)	(6.16)	(0.14)	(6.14)	(−0.06)	(4.17)	(−0.42)	(3.48)
λ^{Δ}	269	−0.06	3.71***	−0.30	8.57***	0.33	2.93***	−1.13	7.18***
		(−0.24)	(4.88)	(−0.54)	(6.14)	(1.32)	(4.10)	(−1.56)	(3.74)
λ^S	100	6.70***	6.96***	1.11	8.61***	7.76**	5.34*	2.05	5.88
		(2.80)	(3.64)	(0.64)	(3.67)	(2.39)	(1.80)	(0.83)	(1.56)
λ^S	269	6.86***	5.70***	0.72	8.47***	8.10**	4.43	1.27	6.38*
		(2.86)	(2.99)	(0.42)	(3.60)	(2.50)	(1.43)	(0.52)	(1.70)
λ^{MF}	–	6.92**	1.99	1.02	−0.10	7.78**	1.50	2.41	−0.79
		(2.83)	(1.01)	(0.58)	(−0.04)	(2.39)	(0.48)	(0.97)	(−0.19)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123
Panel B: Equal-weighted stock portfolios									
λ^{Δ}	100	−0.27	5.15***	2.64***	8.26***	0.31	2.68**	2.22	6.05***
		(−0.58)	(5.92)	(2.65)	(5.68)	(0.51)	(2.26)	(1.61)	(2.99)
λ^{Δ}	269	0.26	4.91***	2.50**	9.86***	1.07	2.03*	1.99	7.69***
		(0.5)	(4.92)	(2.26)	(6.35)	(1.60)	(1.72)	(1.34)	(3.48)
λ^S	100	6.65***	7.14***	3.66*	8.16***	8.09**	4.17	4.62	5.26
		(2.77)	(3.69)	(1.85)	(3.46)	(2.49)	(1.43)	(1.65)	(1.38)
λ^S	269	7.18***	6.90***	3.51*	9.76***	8.85***	3.52	4.40	6.89*
		(3.01)	(3.25)	(1.70)	(4.03)	(2.75)	(1.10)	(1.52)	(1.76)
λ^{MF}	–	6.92***	1.99	1.02	−0.10	7.78**	1.50	2.41	−0.79
		(2.83)	(1.01)	(0.58)	(−0.04)	(2.39)	(0.48)	(0.97)	(−0.19)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123

liver consistent estimates of β and γ, and it does not have the "oracle property" by which the variable selection step identifies the correct model and estimates converge at the optimal rate. By contrast, the Zou (2006) adaptive Lasso has these desirable features, which enables us to construct confidence intervals for cross-sectional slopes as though the first-stage regression were OLS. Adaptive Lasso differs from Lasso in placing higher penalties on parameters with little explanatory power by setting $\omega = |\hat{\beta}|^{-\gamma}$. Our penalization weights use OLS $\hat{\beta}$s (as in Zou, 2006) and a penalty exponent of $\gamma = 1$.

The obvious concern when using Lasso is the selection of the penalization parameter κ. Following standard practice (e.g., Bühlmann and van de Geer, 2011; Hastie et al., 2015), we use k-fold cross-validation to select κ. Cross-validation works as follows. First, select a candidate value of κ_m and partition the sample into k equal "folds'.' We choose the MATLAB default of $k = 10$. Next, for each fold, estimate the model on the set difference of the full sample and the partition. Then, calculate the mean squared error (MSE) of the estimated model on the fold that was set aside. This procedure provides k pseudo-out-of-sample (POOS) MSEs as a function of κ_m. Finally, repeat this procedure for a range of κ_m, and select κ as the value κ_m that maximizes the average POOS MSE. Intuitively, this process tames overfitting by selecting the model with the best out-of-sample predictive properties.[38]

Table A.2 presents results using the adaptive Lasso first stage described by Eq. (D.1). Most importantly, the coeffi-

[38] Remarkably, Chetverikov et al. (2017) demonstrate that time series betas estimated using the cross-validated Lasso converge to the true betas at rate $\sqrt{n}$, up to a negligible log term. Because the convergence rate is comparable to that of OLS, using (adaptive) Lasso in the first stage does not exacerbate the errors-in-variables problem endemic to Fama-MacBeth regressions. We therefore follow standard practice in taking betas as known inputs into the Fama-MacBeth cross-sectional regressions and adjusting standard errors for heteroskedasticity and serial correlation by Newey-West.

Table A.3
Implementation cost estimates for ETFs in Fama-MacBeth regressions.
This table reports Fama-MacBeth estimates of domestic equity exchange traded fund implementation costs (Panel A) and compensation for factor exposure (Panel B). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$,

$$r_{it} = \sum_k \lambda^S_{kt} \hat{\beta}_{ik} 1_{i \in S} + \sum_k \lambda^{ETF}_{kt} \hat{\beta}_{ik} 1_{i \in ETF} + \epsilon_{it}, \ \ t = 1, \ldots, T,$$

where k indexes the four Carhart (1997) factors and λ^Δ is defined as $\lambda^S - \lambda^{ETF}$. The top table uses our baseline specification, and the bottom panel uses the liquidity principal-component augmented specification. Stock portfolio sets are described in Section 3. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		Baseline Specification				Liquidity principal components			
	N_S	*MKT*	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
Panel A: Value-weighted stock portfolios									
λ^Δ	100	0.02	4.60***	−1.91**	−8.67	−0.34	4.47***	−1.70*	−13.48**
		(0.04)	(3.13)	(−2.04)	(−1.49)	(−0.64)	(3.09)	(−1.67)	(−2.14)
λ^Δ	269	0.09	3.91***	−2.40***	−7.27	−0.29	3.89***	−2.11**	−12.26**
		(0.23)	(2.80)	(−2.78)	(−1.27)	(−0.68)	(2.74)	(−2.13)	(−2.03)
λ^{ETF}	–	9.58**	−1.94	4.27**	9.60	10.07**	−1.92	4.01*	14.65*
		(2.36)	(−0.58)	(2.01)	(1.26)	(2.54)	(−0.59)	(1.86)	(1.94)
Panel B: Equal-weighted stock portfolios									
λ^Δ	100	0.15	4.44**	−2.59**	−5.54	−0.46	3.68*	−2.20	−16.77**
		(0.22)	(2.23)	(−2.22)	(−0.92)	(−0.60)	(1.82)	(−1.65)	(−2.27)
λ^Δ	269	0.25	6.06***	−3.82***	−3.26	−0.43	5.39***	−3.23**	−14.77**
		(0.50)	(3.43)	(−3.47)	(−0.61)	(−0.73)	(2.82)	(−2.42)	(−2.26)
λ^{ETF}	–	8.41	−2.92	5.11**	5.26	10.07***	−1.92	4.01*	14.65*
		(1.45)	(−0.64)	(2.02)	(0.55)	(2.54)	(−0.59)	(1.86)	(1.94)

cients on λ^Δ are of similar size and statistical significance as they are in the preceding two tables. Using the adaptive Lasso results in one key change from Table 3, however: the point estimate for *UMD* compensation for mutual funds becomes negligible in the full sample and negative in the recent sample. This feature is consistent with mutual funds earning compensation for momentum exposure only to the extent that momentum also embeds liquidity risk. By including a rich set of liquidity and liquidity risk proxies rather than two principal components, we allow this source of compensation to be spanned in the first stage, thereby effectively kicking out *UMD* as a compensated factor for mutual funds.

Appendix E. Implementation costs for exchange traded products

Our analysis focuses exclusively on implementation costs of mutual funds. The methodology can be applied to other asset classes, as well. In this appendix we consider a relatively new but increasingly important asset class, exchange traded funds. The primary hurdle to implementing this analysis is that ETFs are a relatively new innovation, and factor loadings are not sufficiently diverse to identify cross-sectional slopes until recently. The first US-listed ETF, the Standard & Poors Depositary Receipt (SPDR) S&P 500 (SPY), was introduced in 1993, and value, growth, small-cap, and large-cap ETFs were only introduced by iShares in June 2000. By 2003, there were only 123 US ETFs, with a total AUM of only \$151 billion, about 2% of aggregate mutual fund AUM at the time.[39] Of these, the vast majority were large-cap and market trackers.

The early and mid-2000s saw the addition of ETFs split by market capitalization, style, and sector. Prior to this time, identifying factor premia and implied implementation costs are not possible with (non-)market factor loadings clustered so tightly around (zero) one. In 2008–2009 the Securities and Exchange Commission (SEC) freed ETF providers from tracking pre-defined indexes.[40] This rule change opened the floodgates to the proliferation of ETFs that are evident today. Our sample runs from 2003 to the present to balance sample length and cross-sectional diversity in ETF factor exposures.

We source ETF data from the CRSP mutual fund database. We retain funds that have the “ET_Flag” variable equal to “F” (for exchange-traded funds) or “N” (for exchange-traded notes). We then filter the data as in Appendix A. Table A.3 reports implementation costs and compensation for factor exposure for exchange-traded funds and notes. As was true for mutual funds, exchange-traded funds suffer negligible costs for market factor exposure. Also like mutual funds, the costs for value-factor exposure are high, and they eliminate the return to value. This finding is not to say that value ETFs do not earn a value premium but, rather, that differences in value beta between ETFs are not compensated (or even negatively

[39] See Deutsche Bank Markets Research “ETF Annual Review and Outlook”.
[40] See, e.g., page 14633 of the proposed rule for exchange-traded funds (https://www.sec.gov/rules/proposed/2008/33-8901fr.pdf), also referenced in Chinco and Fos (2018).

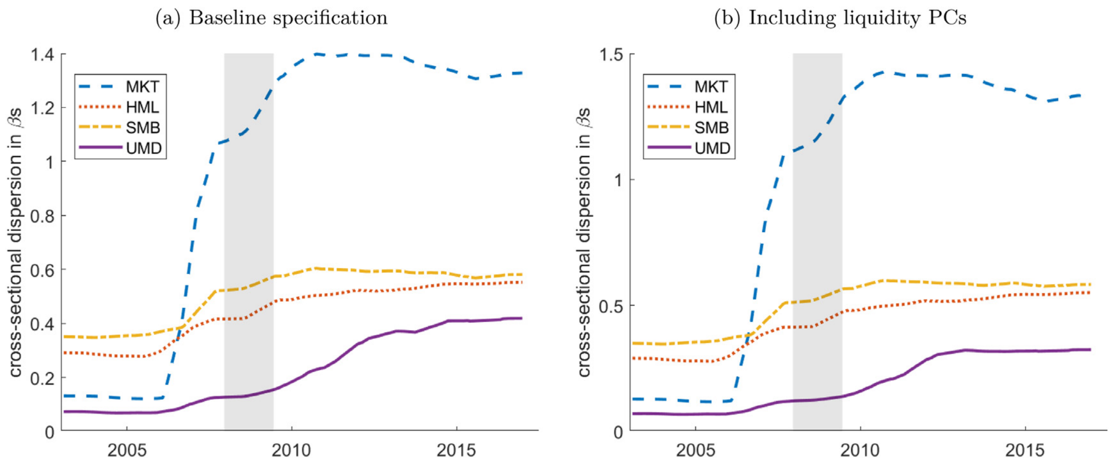

Fig. A.2. Cross-sectional dispersion in exchange traded fund betas over time. This figure plots the cross-sectional dispersion in exchange traded fund betas (for active funds) over time. Betas are estimated using time series regressions for a Carhart (1997) four-factor model. Panel A plots betas in the baseline regression specification and Panel B adds the first principal components of market and funding liquidity proxies. NBER recessions are in gray.

compensated). Size appears to be rewarded in ETFs, even more so than in stocks.

In contrast with mutual funds, ETFs appear to earn a substantial momentum premium. This result, however, is illusory. Point estimates in the baseline, large as they are, are insignificant. The point estimates are even larger in the liquidity-augmented regression, however in both cases these coefficients are not well identified because the cross-sectional variation in momentum betas among ETFs is small, and especially so after netting out the liquidity principal components. Despite the sharp increase in the number of ETFs during this period, not enough momentum and contrarian ETFs exist for a long enough span to recover a meaningful estimate for $\bar{\lambda}^{ETF}_{UMD}$ or its difference from $\bar{\lambda}^{S}_{UMD}$.

Fig. A.2 highlights the limited dispersion in the cross section of momentum betas. The SEC ETF rule change manifests as a marked increase in variation in market betas and a moderate increase in spread of value and size factor betas. Throughout, momentum factor betas exhibit far less variation, especially when liquidity principal components are included. While we do not have a formal criterion for sufficient dispersion to identify λ^{ETF}_{UMD}, it is apparent that we do not have enough for our estimates for momentum compensation or costs to be reliable.

Appendix F. Errors-in-variables and other biases in estimated mutual fund factor compensation

Section 4.1 motivates the use of static betas to resolve the errors-in-variables problem in Fama-MacBeth regression. In this appendix, we conduct three analyses to verify that this approach succeeds. First, as suggested by a referee, we report the average standard errors for betas across the stock portfolio and mutual fund samples. Within each date, we compute the average and median standard error for stock and mutual fund portfolios. We then average these values across dates. Table A.4 reports these values. Mutual fund betas are typically measured with less precision than stock portfolio betas, but differences in attenuation are usually mild, in the 3%–5% range. Momentum coefficients are somewhat more attenuated for mutual funds, but even there, differences in attenuation can explain at most (roughly) a 10% reduction in the estimated compensation to momentum. To illustrate, in our baseline specification in Table 2, we find $\hat{\lambda}^{S}_{UMD} = 8.85$ and $\hat{\lambda}^{MF}_{UMD} = 1.54$, leading to $\hat{\lambda}^{\Delta}_{UMD} = 7.31$. If we use the attenuation factors in Table A.4 to “undo” the bias in these estimates, $\tilde{\lambda}^{S}_{UMD} = \frac{1}{1-0.02} 8.85 = 9.03$ and $\tilde{\lambda}^{MF}_{UMD} = \frac{1}{1-0.14} 1.54 = 1.79$. Consequently, $\tilde{\lambda}^{\Delta}_{UMD} = \tilde{\lambda}^{S}_{UMD} - \tilde{\lambda}^{MF}_{UMD} = 7.24$, which is only marginally smaller than our initial estimate of $\hat{\lambda}^{\Delta}_{UMD} = 7.31$.

Second, to address concerns about attenuation or bias in estimated mutual fund compensation, we run regressions allowing for nonzero mutual fund alphas. To do this, we reestimate our mutual fund compensation regressions with a constant term in the cross-sectional regression step for the mutual fund portfolios. We tabulate values for our baseline specification and the principal-component augmented specification in Table A.5.

Comparing mutual fund compensation estimates with and without an included constant, we observe that all λ_{MF} coefficients are virtually the same. The one exception is that full-sample (second-half) compensation for market risk is about 20 basis points (bps) higher (60 bps smaller) on a base of 7%–8% in the baseline specification. No evidence exists of meaningful bias in mutual fund compensation associated with forcing expected returns through the origin.

Third, we use the new instrumental variable technique of Jegadeesh et al. (2019) to undo the effects of measurement error in our generated regressors. As a first step, we estimate betas twice: once using only data from even-numbered months, and once using only data from odd-numbered months. We then instrument even-month betas with the preceding odd-month betas, and vice versa. For example, for odd months, this step consists of regressing

$$\beta^{(odd)}_{ik} = a_t + \sum_k b_{kt} \beta^{(even)}_{ik} + \epsilon_{ikt} \tag{F.1}$$

Table A.4
Average and median standard errors of time series betas for stock and mutual fund portfolios.
This table reports averages of the mean and median standard errors of betas within each cross-section. The estimated attenuation factor is the percent reduction in cross-sectional slope estimates assuming all other factor betas are measured perfectly, that is, $A = 1 - 1/\left(1 + \overline{SE_k^2}/var\left(\hat{\beta}_k\right)\right)$. Time series betas are estimated using the four Carhart (1997) factors. Stock portfolio sets are described in Section 3, and we use the more expansive $N_S = 269$ portfolio set.

	1970 – 2016				1993 – 2016			
	MKT	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
Panel A: Baseline specification								
Mean SE$\left(\beta_k^{MF}\right)$	0.03	0.05	0.05	0.03	0.04	0.05	0.05	0.03
Mean SE$\left(\beta_k^{S}\right)$	0.02	0.03	0.03	0.02	0.03	0.05	0.04	0.03
Median SE$\left(\beta_k^{MF}\right)$	0.03	0.04	0.04	0.02	0.03	0.04	0.04	0.03
Median SE$\left(\beta_k^{S}\right)$	0.02	0.03	0.02	0.02	0.03	0.04	0.03	0.02
Attenuation (MF)	0.04	0.06	0.05	0.14	0.03	0.04	0.04	0.12
Attenuation (S)	0.04	0.02	0.01	0.02	0.05	0.03	0.01	0.03
Panel B: Including liquidity principal components								
Mean SE$\left(\beta_k^{MF}\right)$	0.04	0.05	0.05	0.03	0.04	0.06	0.05	0.03
Mean SE$\left(\beta_k^{S}\right)$	0.02	0.03	0.03	0.02	0.04	0.05	0.04	0.03
Median SE$\left(\beta_k^{MF}\right)$	0.03	0.04	0.04	0.03	0.03	0.05	0.04	0.03
Median SE$\left(\beta_k^{S}\right)$	0.02	0.03	0.02	0.02	0.03	0.04	0.03	0.02
Attenuation (MF)	0.04	0.07	0.05	0.14	0.03	0.06	0.04	0.14
Attenuation (S)	0.04	0.03	0.01	0.02	0.06	0.04	0.01	0.04

Table A.5
Mutual fund compensation estimates with and without a constant.
This table reports Fama-MacBeth estimates of the compensation for factor exposure for domestic equity mutual funds. Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$,

$$r_{it} = \alpha_t^{MF} 1_{i \in MF} + \sum_k \lambda_{kt}^{MF} \hat{\beta}_{ik} 1_{i \in MF} + \epsilon_{it}, \quad t = 1, \ldots, T,$$

where k indexes the four Carhart (1997) factors. First-stage regression estimates in both tables include these factors, and the second table also includes the first principal component of market liquidity proxies and the first principal component of funding liquidity proxies. Liquidity proxies and stock portfolio sets are described in Section 3. The first row allows α_t^{MF} to be estimated freely and the second row sets α_t^{MF} to be zero, as in the main text. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1970 – 2016				1993 – 2016			
	α_t	*MKT*	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
Panel A: Baseline specification									
λ^{MF}	$\neq 0$	7.18***	2.64	1.02	1.47	7.24**	2.18	2.27	1.67
		(2.94)	(1.56)	(0.60)	(0.60)	(2.26)	(0.80)	(0.95)	(0.43)
λ^{MF}	$= 0$	6.98***	2.62	1.01	1.54	7.78**	2.31	2.20	1.73
		(2.86)	(1.51)	(0.59)	(0.63)	(2.38)	(0.83)	(0.92)	(0.45)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123
Panel B: Including liquidity principal components									
λ^{MF}	$\neq 0$	6.99***	2.64	0.90	1.28	7.80**	2.09	2.22	0.76
		(2.87)	(1.51)	(0.53)	(0.52)	(2.41)	(0.74)	(0.92)	(0.20)
λ^{MF}	$= 0$	6.98***	2.62	1.01	1.54	7.78**	2.31	2.20	1.73
		(2.86)	(1.51)	(0.59)	(0.63)	(2.38)	(0.83)	(0.92)	(0.45)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123

Table A.6
Implementation cost estimates, baseline specification with instrumental variables.
This table reports instrumented Fama-MacBeth estimates of the compensation for factor exposure for stock portfolios (third and fourth rows), domestic equity mutual funds (fifth row), and their difference (first two rows). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$,

$$r_{it} = \sum_k \lambda^S_{kt} \hat{\beta}_{ik} 1_{i \in S} + \sum_k \lambda^{MF}_{kt} \hat{\beta}_{ik} 1_{i \in MF} + \epsilon_{it}, \ t = 1, \ldots, T,$$

where k indexes the four Carhart (1997) factors, λ^Δ is defined as $\lambda^S - \lambda^{MF}$, and $\hat{\beta}$s are estimated using alternating months as in Jegadeesh et al. (2019). Stock portfolio sets are described in Section III. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1970 – 2016				1993 – 2016			
	N_S	*MKT*	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
Panel A: Value-weighted stock portfolios									
λ^Δ	100	-0.43	4.79***	0.00	8.68***	0.03	3.77***	-0.51	6.45***
		(-1.42)	(5.24)	(0.00)	(4.58)	(0.10)	(3.71)	(-0.58)	(2.62)
λ^Δ	269	-0.30	3.48***	-0.33	8.54***	0.28	2.44***	-0.94	7.04***
		(-1.24)	(4.09)	(-0.60)	(4.62)	(1.21)	(3.27)	(-1.27)	(3.00)
λ^S	100	6.60***	7.03***	1.04	9.36***	7.83**	5.72*	1.65	7.50*
		(2.74)	(3.64)	(0.60)	(3.66)	(2.40)	(1.89)	(0.66)	(1.74)
λ^S	269	6.72***	5.72***	0.70	9.22***	8.08**	4.38	1.21	8.09**
		(2.80)	(3.05)	(0.41)	(3.90)	(2.49)	(1.47)	(0.50)	(2.15)
λ^{MF}	–	7.02***	2.24	1.03	0.68	7.80**	1.95	2.15	1.05
		(2.87)	(1.23)	(0.60)	(0.25)	(2.38)	(0.68)	(0.89)	(0.25)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123
Panel B: Equal-weighted stock portfolios									
λ^Δ	100	-0.27	5.64***	1.93*	9.24***	0.31	4.47***	1.86	7.11***
		(-0.55)	(5.79)	(1.94)	(4.80)	(0.49)	(3.96)	(1.29)	(2.79)
λ^Δ	269	0.31	4.50***	1.61	10.37***	1.10	3.18***	1.66	9.19***
		(0.59)	(4.15)	(1.45)	(5.45)	(1.63)	(2.71)	(1.04)	(3.64)
λ^S	100	6.76***	7.88***	2.96	9.92***	8.10**	6.41**	4.02	8.16*
		(2.81)	(4.14)	(1.48)	(3.93)	(2.49)	(2.17)	(1.37)	(1.91)
λ^S	269	7.33***	6.74***	2.65	11.06***	8.90***	5.13*	3.81	10.23***
		(3.05)	(3.35)	(1.26)	(4.51)	(2.72)	(1.67)	(1.27)	(2.58)
λ^{MF}	–	7.02***	2.24	1.03	0.68	7.80**	1.95	2.15	1.05
		(2.87)	(1.23)	(0.60)	(0.25)	(2.38)	(0.68)	(0.89)	(0.25)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1286	1286	1286	1286	2123	2123	2123	2123

for each factor k. We reestimate this specification for each date t to account for potential time-variation in attenuation associated with changes in the cross section of active mutual funds. Funds that do not exist in both months t and $t-1$ are dropped. Jegadeesh et al. (2019) demonstrate in a multifactor setting that once *predicted* betas are used in place of full-sample betas, cross-sectional slopes are no longer attenuated and the Fama-MacBeth procedure and standard error construction otherwise can be applied unchanged.

Table A.6 reports instrumented estimates of implementation costs. Compared with Table 2, the point estimate for mutual funds' compensation is smaller for all four factors in the instrumented analysis. Implied implementation costs rise as a consequence for *HML* and *UMD*, although the greater noise in cross-sectional slopes leads to an ambiguous net effect on the statistical significance of these compensation differentials. In short, as before, we find that implementation costs eliminate real-world compensation to value and momentum factors and have no effect on compensation for market and size factors.

Appendix G. Compensation and costs of factor-targeting mutual funds

In this appendix we consider the implementation costs of mutual funds that specifically target factors. We do so because the compensation to incidental factor exposures incurred as part of other strategies can differ from the compensation accruing to funds whose primary objective is to harvest particular factor premia.

Evaluating the costs of targeting funds faces two challenges. The first is how to identify whether funds target a factor. While funds that exclusively target the market are readily identified by their high R^2s in time series regressions of their returns on the market return, funds that target other factors with different implementations from the academic one have lower R^2s, and Lettau et al. (2018) note that "even funds with an explicit 'value' objective hold a larger share of low [book-to-market ratio] stocks than high-BM stocks." Similarly, self-designated benchmarks are often wrong, and perhaps intentionally so to give rise to higher apparent relative performance (Sensoy, 2009).

Table A.7
Implementation cost estimates, baseline specification with $|t| \geq 2$.
This table reports Fama-MacBeth estimates of the compensation for factor exposure for stock portfolios (third and fourth rows), domestic equity mutual funds (fifth row), and their difference (first two rows). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$,

$$r_{it} = \sum_k \lambda_{kt}^S \hat{\beta}_{ik} 1_{i \in S} + \sum_k \lambda_{kt}^{MF} \hat{\beta}_{ik} 1_{i \in MF} + \epsilon_{it}, \ \ t = 1, \ldots, T,$$

where k indexes the four Carhart (1997) factors and λ^Δ is defined as $\lambda^S - \lambda^{MF}$. We run K separate regressions using only mutual funds with $|t_k| \geq 2$ and report average compensation and compensation differentials in the corresponding column. Stock portfolio sets are described in Section 3. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1970 – 2016				1993 – 2016			
	N_S	*MKT*	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
Panel A: Value-weighted stock portfolios									
λ^Δ	100	-0.36	3.61***	0.34	6.58***	-0.11	3.35***	0.03	4.16**
		(-1.21)	(4.53)	(0.56)	(5.10)	(-0.32)	(4.04)	(0.03)	(2.54)
λ^Δ	269	-0.19	2.38***	0.01	6.71***	0.28	2.32***	-0.70	4.93***
		(-0.79)	(3.36)	(0.02)	(5.01)	(1.25)	(3.66)	(-1.03)	(2.78)
λ^S	100	6.60***	6.43***	1.27	8.72***	7.67**	5.43*	1.96	6.01
		(2.75)	(3.51)	(0.75)	(3.74)	(2.35)	(1.93)	(0.81)	(1.60)
λ^S	269	6.77***	5.20***	0.94	8.85***	8.06**	4.40	1.23	6.78*
		(2.82)	(2.84)	(0.56)	(3.80)	(2.49)	(1.54)	(0.51)	(1.83)
λ^{MF}	–	6.96***	2.82	0.93	2.14	7.78**	2.08	1.93	1.84
		(2.85)	(1.63)	(0.53)	(0.88)	(2.38)	(0.75)	(0.81)	(0.48)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1279	941	919	711	2106	1580	1520	1131
Panel B: Equal-weighted stock portfolios									
λ^Δ	100	-0.34	4.27***	2.42**	6.23***	0.07	3.40***	2.41*	3.61**
		(-0.72)	(5.03)	(2.54)	(4.81)	(0.12)	(3.47)	(1.77)	(2.12)
λ^Δ	269	0.27	3.11***	2.30**	7.92***	0.95	2.25**	2.31	5.93***
		(0.54)	(3.27)	(2.17)	(5.67)	(1.45)	(2.17)	(1.54)	(3.03)
λ^S	100	6.62***	7.09***	3.35***	8.37***	7.85**	5.48**	4.34	5.45
		(2.75)	(3.91)	(1.70)	(3.59)	(2.39)	(1.99)	(1.53)	(1.44)
λ^S	269	7.23***	5.93***	3.23	10.06***	8.73***	4.33	4.25	7.78**
		(3.02)	(3.03)	(1.56)	(4.17)	(2.69)	(1.47)	(1.43)	(1.98)
λ^{MF}	–	6.96***	2.82	0.93	2.14	7.78**	2.08	1.93	1.84
		(2.85)	(1.63)	(0.53)	(0.88)	(2.38)	(0.75)	(0.81)	(0.48)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1279	941	919	711	2106	1580	1520	1131

Funds can target multiple factors or none at all. For these reasons, we take a regression-based approach to assessing whether funds target factors as in Sharpe (1992) and Fung and Hsieh (1997). We denote a fund as targeting a factor if the t-statistic of its beta with respect to a factor exceeds two in absolute value. We use the t-statistic to capture whether a fund reliably loads on a factor in a long or short direction, that is, whether it has more factor exposure than would be expected by chance.

The second challenge to focusing on funds that target a factor is that different funds target different factors. Running a multivariate regression using only the funds that target a factor identifies factor compensation well for the targeted factor but not necessarily for the other factors. Nevertheless, a multivariate specification is necessary to clean up residual exposures to non-targeted factors that may otherwise explain differences in compensation. For this reason, we run separate regressions for the subsamples of funds that target *MKT, HML, SMB*, and *UMD*, and we retain coefficients only for the targeted factor. For example, only 55% of funds in the 1970–2016 have momentum betas statistically distinguishable from zero, and we tabulate the estimated momentum compensation only for this group of funds (and comparable stock portfolios).

With these adjustments made, Tables A.7 and A.8 report implementation costs and factor compensations for targeting mutual funds. Relative to the baseline factor compensation estimates, factor compensation for this set of funds is about 20 bps larger for *HML* and up to 60 bps larger for *UMD* in Table A.7; including liquidity principal components in Table A.8 has minimal incremental effect. The resulting implementation costs remain economically large and correspond with significant or near-complete attenuation of factor premia, respectively. Statistical significance drops slightly for non-market factors because sample sizes shrink when restricting only to targeting mutual funds and, consequently, realized factor premia are estimated with less precision. In short, we see little difference

Table A.8
Implementation cost estimates, liquidity principal components with $|t| \geq 2$.
This table reports Fama-MacBeth estimates of the compensation for factor exposure for stock portfolios (third and fourth rows), domestic equity mutual funds (fifth row), and their difference (first two rows). Coefficients are the average cross-sectional slopes $\bar{\lambda}_k$ across monthly regressions of excess returns r_{it} on time series betas $\hat{\beta}_{ik}$,

$$r_{it} = \sum_k \lambda_{kt}^S \hat{\beta}_{ik} 1_{i\in S} + \sum_k \lambda_{kt}^{MF} \hat{\beta}_{ik} 1_{i\in MF} + \epsilon_{it}, \ t = 1, \ldots, T,$$

where k indexes the four Carhart (1997) factors and λ^Δ is defined as $\lambda^S - \lambda^{MF}$. We run K separate regressions using only mutual funds with $|t_k| \geq 2$ and report average compensation and compensation differentials in the corresponding column. First-stage regression estimates include these factors, the first principal component of market liquidity proxies, and the first principal component of funding liquidity proxies. Liquidity proxies and stock portfolio sets are described in Section 3. All coefficients are annualized and reported in percent. Standard errors are Newey-West with three lags. t statistics are reported in parentheses. Parameters different from zero at the 10%, 5% or 1% significance levels are marked with one, two or three asterisks.

		1970 – 2016				1993 – 2016			
	N_S	*MKT*	*HML*	*SMB*	*UMD*	*MKT*	*HML*	*SMB*	*UMD*
Panel A: Value-weighted stock portfolios									
λ^Δ	100	−0.42	3.73***	0.29	7.99***	−0.12	3.50***	−0.04	5.87***
		(−1.38)	(4.50)	(0.47)	(6.19)	(−0.37)	(4.05)	(−0.05)	(3.48)
λ^Δ	269	−0.20	2.48***	−0.08	8.05***	0.27	2.50***	−0.73	6.48***
		(−0.83)	(3.31)	(−0.15)	(6.10)	(1.21)	(3.86)	(−1.05)	(3.64)
λ^S	100	6.55***	6.71***	1.26	8.77***	7.68**	5.38*	1.98	5.99
		(2.74)	(3.63)	(0.74)	(3.76)	(2.37)	(1.90)	(0.82)	(1.59)
λ^S	269	6.77***	5.47***	0.89	8.84***	8.08**	4.39	1.30	6.60*
		(2.83)	(2.94)	(0.53)	(3.78)	(2.51)	(1.51)	(0.54)	(1.78)
λ^{MF}	–	6.97***	2.98*	0.97	0.78	7.81**	1.89	2.02	0.12
		(2.86)	(1.70)	(0.56)	(0.33)	(2.41)	(0.67)	(0.83)	(0.03)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1279	908	914	709	2104	1527	1505	1118
Panel B: Equal-weighted stock portfolios									
λ^Δ	100	−0.51	4.14***	2.63***	7.42***	0.00	2.88***	2.51*	4.65***
		(−1.08)	(4.73)	(2.76)	(5.62)	(0.00)	(2.73)	(1.88)	(2.65)
λ^Δ	269	0.07	3.30***	2.58**	9.03***	0.74	2.13**	2.49*	6.37***
		(0.13)	(3.36)	(2.45)	(6.28)	(1.09)	(1.99)	(1.73)	(3.14)
λ^S	100	6.46***	7.12***	3.60*	8.20***	7.81**	4.76*	4.54	4.76
		(2.70)	(3.88)	(1.84)	(3.51)	(2.40)	(1.73)	(1.63)	(1.26)
λ^S	269	7.04***	6.28***	3.55*	9.82***	8.55***	4.01	4.51	6.49
		(2.97)	(3.14)	(1.73)	(4.05)	(2.66)	(1.35)	(1.57)	(1.64)
λ^{MF}	–	6.97***	2.98*	0.97	0.78	7.81**	1.89	2.02	0.12
		(2.86)	(1.70)	(0.56)	(0.33)	(2.41)	(0.67)	(0.83)	(0.03)
T		564	564	564	564	282	282	282	282
$\bar{N}_{MF}$		1279	908	914	709	2104	1527	1505	1118

in costs between factor-targeting mutual funds and the full set of mutual funds.

References

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. J. Financ. Mark. 5 (1), 31–56.
Anand, A., Irvine, P., Puckett, A., Venkataraman, K., 2011. Performance of institutional trading desks: an analysis of persistence in trading costs. Rev. Financ. Stud. 25 (2), 557–598.
Arnott, R.D., Kalesnik, V., Wu, L.J., 2017. The incredible shrinking factor return. Working Paper
Asness, C.S., Moskowitz, T.J., Pedersen, L.H., 2013. Value and momentum everywhere. J. Finance 68 (3), 929–985.
Banz, R.W., 1981. The relationship between return and market value of common stocks. J. Financ. Econ. 9 (1), 3–18.
Barber, B.M., Huang, X., Odean, T., 2016. Which factors matter to investors? Evidence from mutual fund flows. Rev. Financ. Stud. 29 (10), 2600–2642.
Barroso, P., Santa-Clara, P., 2015. Momentum has its moments. J. Financ. Econ. 116 (1), 111–120.
Basak, S., Pavlova, A., 2013. Asset prices and institutional investors. Am. Econ. Rev. 103 (5), 1728–1758.
Basu, S., 1977. Investment performance of common stocks in relation to their price-earnings ratios: a test of the efficient market hypothesis. J. Finance 32 (3), 663–682.
Berk, J.B., 2000. Sorting out sorts. J. Finance 55 (1), 407–427.
Berk, J.B., van Binsbergen, J.H., 2015. Measuring skill in the mutual fund industry. J. Financ. Econ. 118 (1), 1–20.
Berk, J.B., van Binsbergen, J.H., 2016. Assessing asset pricing models using revealed preference. J. Financ. Econ. 119 (1), 1–23.
Berk, J.B., Green, R.C., 2004. Mutual fund flows and performance in rational markets. J. Polit. Econ. 112 (6), 1269–1295.
Boyson, N.M., Stahel, C.W., Stulz, R.M., 2010. Hedge fund contagion and liquidity shocks. J. Finance 65 (5), 1789–1816.
Breen, W.J., Hodrick, L.S., Korajczyk, R.A., 2002. Predicting equity liquidity. Manage. Sci. 48 (4), 470–483.
Brunnermeier, M.K., 2009. Deciphering the liquidity and credit crunch 2007–2008. J. Econ. Perspect. 23 (1), 77–100.
Bühlmann, P., van de Geer, S., 2011. Statistics for High-Dimensional Data: Methods, Theory, and Applications. Springer Series in Statistics.
Carhart, M.M., 1997. On persistence in mutual fund performance. J. Finance 52 (1), 57–82.
Chen, Z., Stanzl, W., Watanabe, M., 2002. Price impact costs and the limit of arbitrage. Working Paper
Chetverikov, D., Liao, Z., Chernozhukov, V., 2017. On cross-validated lasso. Working Paper

Chinco, A., Fos, V., 2018. The sound of many funds rebalancing. Working Paper
Cochrane, J.H., 2011. Presidential address: discount rates. J. Finance 66 (4), 1047–1108.
Corwin, S.A., Schultz, P., 2012. A simple way to estimate bid-ask spreads from daily high and low prices. J. Finance 67 (2), 719–760.
Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. J. Finance 52 (3), 1035–1058.
Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross sectional variation in stock returns. J. Finance 52 (1), 1–33.
Davis, J.L., Fama, E.F., French, K.R., 2000. Characteristics, covariances, and average returns: 1929 to 1997. J. Finance 55 (1), 389–406.
Edelen, R.M., Evans, R.B., Kadlec, G.B., 2007. Scale effects in mutual fund performance: the role of trading costs. Working Paper
Edelen, R.M., Evans, R.B., Kadlec, G.B., 2013. Shedding light on "invisible" costs: trading costs and mutual fund performance. Financ. Anal. J. 69 (1), 33–44.
Elton, E.J., Gruber, M.J., Blake, C.R., 2001. A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and morningstar mutual fund databases. J. Finance 56 (6), 2415–2430.
Engle, R., Ferstenberg, R., Russell, J., 2012. Measuring and modeling execution cost and risk. J. Portf. Manage. 38 (2), 14–28.
Fama, E.F., 1970. Efficient capital markets: a review of theory and empirical work. J. Finance 25 (2), 383–417.
Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. J. Finance 47 (2), 427–465.
Fama, E.F., French, K.R., 2008. Dissecting anomalies. J. Finance 63 (4), 1653–1678.
Fama, E.F., French, K.R., 2010. Luck versus skill in the cross-section of mutual fund returns. J. Finance 65 (5), 1915–1947.
Fama, E.F., French, K.R., 2012. Size, value, and momentum in international stock returns. J. Financ. Econ. 105 (3), 457–472.
Fama, E.F., French, K.R., 2018. Comparing cross-section and time-series factor models. Working Paper
Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. J. Polit. Econ. 81 (3), 607–636.
Frazzini, A., Israel, R., Moskowitz, T.J., 2015. Trading costs of asset pricing anomalies. Working Paper
Frazzini, A., Pedersen, L.H., 2014. Betting against beta. J. Financ. Econ. 111 (1), 1–25.
Fung, W., Hsieh, D.A., 1997. Empirical characteristics of dynamic trading strategies: the case of hedge funds. Rev. Financ. Stud. 10 (2), 275–302.
Glosten, L.R., Harris, L.E., 1988. Estimating the components of the bid/ask spread. J. Financ. Econ. 21 (1), 123–142.
Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior. Am. Econ. Rev. 85 (5), 1088–1105.
Grundy, B.D., Martin, J.S., 2001. Understanding the nature of the risks and the source of the rewards to momentum investing. Rev. Financ. Stud. 14 (1), 29–78.
Harvey, C.R., 2017. Presidential address: the scientific outlook in financial economics. J. Finance 72 (4), 1399–1440.
Harvey, C.R., Liu, Y., Zhu, H., 2016. ... and the cross-section of expected returns. Rev. Financ. Stud. 29 (1), 5–68.
Hasbrouck, J., 2009. Trading costs and returns for U.S. equities: estimating effective costs from daily data. J. Finance 64 (3), 1445–1477.
Hastie, T., Tibshirani, R., Wainwright, M., 2015. Statistical Learning with Sparsity: The Lasso and Generalizations. CRC Press.
He, Z., Kelly, B., Manela, A., 2017. Intermediary asset pricing: new evidence from many asset classes. J. Financ. Econ..
Hou, K., Xue, C., Zhang, L., 2017. Replicating anomalies. Working Paper
Israel, R., Moskowitz, T.J., 2013. The role of shorting, firm size, and time on market anomalies. J. Financ. Econ. 108 (2), 275–301.
Jegadeesh, N., Noh, J., Pukthuanthong, K., Roll, R., Wang, J., 2019. Empirical tests of asset pricing models with individual assets: resolving the errors-in-variables bias in risk premium estimation. J. Financ. Econ..
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. J. Finance 48 (1), 65–91.
Jegadeesh, N., Titman, S., 2001. Profitability of momentum strategies: an evaluation of alternative explanations. J. Finance 56 (2), 699–720.
Jones, C.M., 2002. A century of stock market liquidity and trading costs. Working Paper
Keim, D.B., 2003. The cost of trend chasing and the illusion of momentum profits. Working Paper
Keim, D.B., Madhavan, A., 1997. Transactions costs and investment style: an inter-exchange analysis of institutional equity trades. J. Financ. Econ. 46 (3), 265–292.
Kelly, B.T., Pruitt, S., Su, Y., 2018. Characteristics are covariances: a unified model of risk and return. Working Paper
Korajczyk, R.A., Sadka, R., 2004. Are momentum profits robust to trading costs? J. Finance 59 (3), 1039–1082.
Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica 53 (6), 1315–1335.
Lesmond, D.A., Ogden, J.P., Trzcinka, C.A., 1999. A new estimate of transaction costs. Rev. Financ. Stud. 12 (5), 1113–1141.
Lesmond, D.A., Schill, M.J., Zhou, C., 2004. The illusory nature of momentum profits. J. Financ. Econ. 71 (2), 349–380.
Lettau, M., Ludvigson, S., Manoel, P., 2018. Characteristics of mutual fund portfolios: where are the value funds?. Working Paper
Lettau, M., Maggiori, M., Weber, M., 2014. Conditional risk premia in currency markets and other asset classes. J. Financ. Econ. 114 (2), 197–225.
Lewellen, J., Nagel, S., Shanken, J., 2010. A skeptical appraisal of asset pricing tests. J. Financ. Econ. 96 (2), 175–194.
Nagel, S., 2012. Evaporating liquidity. Rev. Financ. Stud. 25 (7), 2005–2039.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.
Novy-Marx, R., Velikov, M., 2016. A taxonomy of anomalies and their trading costs. Rev. Financ. Stud. 29 (1), 104–147.
Oster, E., 2019. Unobservable selection and coefficient stability: theory and evidence. J. Bus. Econ. Stat. 37 (2), 187–204.
Pastor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. J. Polit. Econ. 111 (3), 642–685.
Pastor, L., Stambaugh, R.F., 2012. On the size of the active management industry. J. Polit. Econ. 120 (4), 740–781.
Pastor, L., Stambaugh, R.F., Taylor, L.A., 2015. Scale and skill in active management. J. Financ. Econ. 116 (1), 23–45.
Patton, A.J., Timmermann, A., 2010. Monotonicity in asset returns: new tests with applications to the term structure, the capm, and portfolio sorts. J. Financ. Econ. 98 (3), 605–625.
Perold, A.F., 1988. The implementation shortfall: paper versus reality. J. Portf. Manage. 14 (3), 4–9.
Roweis, S., 1998. EM algorithms for PCA and SPCA. In: Proceedings of the 1997 Conference on Advances in Neural Information Processing Systems 10. MIT Press, Cambridge, MA, USA, pp. 626–632.
Sadka, R., 2010. Liquidity risk and the cross-section of hedge-fund returns. J. Financ. Econ. 98 (1), 54–71.
Sensoy, B.A., 2009. Performance evaluation and self-designated benchmark indexes in the mutual fund industry. J. Financ. Econ. 92 (1), 25–39.
Shanken, J., 1992. On the estimation of beta-pricing models. Rev. Financ. Stud. 5 (1), 1–55.
Sharpe, W.F., 1992. Asset allocation. J. Portf. Manage. 18 (2), 7–19.
Stock, J.H., Watson, M.W., 2002. Forecasting using principal components from a large number of predictors. J. Am. Stat. Assoc. 97 (460), 1167–1179.
Tibshirani, R., 1994. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B 58, 267–288.
Wermers, R., 2000. Mutual fund performance: an empirical decomposition into stock-picking talent, style, transactions costs, and expenses. J. Finance 55 (4), 1655–1695.
Zou, H., 2006. The adaptive lasso and its oracle properties. J. Am. Stat. Assoc. 101 (476), 1418–1429.