Streaks in Daily Returns*

Alexander Klos[△] Alexandra Koehl[◊] Simon Rottke[□]

June 2020

Abstract

A simple return extrapolation model suggests that streaks in returns, which we define as n-day consecutive over-/under-performance relative to the market, predict future returns. We test this prediction using daily data from international equity markets and find strong empirical support. US-based value-weighted trading strategies have annualized Sharpe ratios around 2 (depending on the specification). We replicate the results in international markets and are able to increase the Sharpe ratio to above 3 by diversifying across regions. We argue that liquidity is unlikely to explain the results as streak portfolio returns based on mid-quote-prices are strongest among stocks with the lowest bid-ask spreads.

Keywords: Streaks, return extrapolation, investor behavior.

JEL-Classification: G12, G4

*We thank Alexander Hillert, Vesa Pursiainen, and conference participants at the SAFE asset pricing workshop and the German Finance association for helpful comments. Part of this paper was written while Simon Rottke was visiting Columbia Business School.

[△]alexander.klos@qber.uni-kiel.de, Kiel University.

[◊]alexandra.koehl@qber.uni-kiel.de, Kiel University.

[□]simon.rottke@uva.nl, University of Amsterdam.

1 Introduction

Theories of return extrapolation have recently received growing attention in the finance literature (Greenwood and Shleifer, 2014, Barberis, Greenwood, Jin, and Shleifer, 2015, 2018, Da, Huang, and Jin, 2019). Consistent with the literature on extrapolation in other fields of economics (see, e.g., Coibion and Gorodnichenko, 2015, Mankiw and Reis, 2002), the time frame over which return extrapolation is analyzed empirically is often a quarter (Barberis, Greenwood, Jin, and Shleifer, 2018). Recent research has looked at shorter horizons. Da, Huang, and Jin (2019) analyze weekly data and find evidence for short-term return extrapolation. We pursue this avenue of investigation and explore the idea that return extrapolation also plays a role in forming daily expectations and consequently impacts prices.

The intuition behind return predictability arising from return extrapolation is the following: A group of extrapolators develops extreme optimistic (pessimistic) sentiment after observing a series of positive (negative) daily returns. If arbitrageurs do not sufficiently trade against the extrapolators' price pressure, the extreme sentiment temporarily finds its way into equilibrium prices. The resulting over- or underpricing tends to reverse on the following trading day. We use a simplified version of the model by Da, Huang, and Jin (2019) to show that streaks in daily returns are a useful proxy to identify stocks with high or low sentiment among extrapolators. In our definition, a streak in daily returns of length n occurs when an asset has outperformed or underperformed the market over n consecutive days.

Our first empirical test of this theoretical result uses daily US stock returns from January 1998 to December 2019. We find that stocks with streaks in daily returns up to day $t-1$ exhibit significant reversals on day t, consistent with the theoretical prediction. Value-weighted long-short portfolios that go long stocks with negative streaks and short stocks with positive streaks earn economically sizeable returns over the next trading day. Returns of long-short portfolios increase monotonically with streak length.

We then move to international data and find similar evidence in equity markets around the globe. Regional streak strategy returns are only modestly correlated with each other. A global diversified streak strategy portfolio therefore increases the Sharpe ratio further.

Subsequently, we revisit the US evidence where data availability allows deeper empirical analyses. We argue that streaks are distinct from classic short-term reversal strategies and are unlikely to be explained by liquidity. Excess returns calculated based on mid-quote prices are still 10.9 bps per day in the US (as opposed to 13.3 bps per day based on closing prices). Portfolio returns based on mid-quotes are largest among the stocks with the smallest bid-ask spreads.

Streak strategy returns cannot be explained by the five Fama-French factors, the short-term reversal factor, the IVOL effect, weekly momentum or a factor constructed based on bid-ask-spread sorts. We additionally match stocks in streak portfolios with control stocks that have similar returns on day $t-1$ but no streak in returns. A value-weighted portfolio of stocks with streaks outperforms a value-weighted portfolio of control stocks on day t and the outperformance increases in streak length. We also regress daily returns on streak dummies, previous day's and cumulative previous n-day's excess returns in market-value-weighted Fama-MacBeth regressions (Fama and MacBeth, 1973, Green, Hand, and Zhang, 2017). Streak coefficients remain statistically significant and economically sizable, highlighting that streaks are distinct from predictability arising from past returns alone.

Returns from streak portfolios are higher on earnings announcement days, consistent with the idea that streaks capture mispricing that gets resolved as new information becomes available. After bid-ask spreads have fallen dramatically in the early 2000s, a *long-only* streak strategy with trading-cost mitigation enhancements outperforms the market after accounting for the costs arising from bid-ask spreads.[1]

[1] We are unaware of previous research investigating the short-term return predictability of streaks in daily returns. Early research considers daily returns after large price declines (defined as price declines of 10 percent or more, Bremer and Sweeney, 1991, Cox and Peterson, 1994), or the stock price performance of the three most extreme daily winners and losers on the subsequent day (Atkins and Dyl, 1990). Theses papers typically rely on small samples, presumably due to the limited ability of computers to deal with large data sets at the time of the papers' writing. A large and diverse literature followed these early papers, performing

2 Inferring New Testable Implications from a Model of Return Extrapolation

We explore the model of Da, Huang, and Jin (2019) to develop previously unstated implications of recent theories on return extrapolation. In Da, Huang, and Jin (2019), agents trade a risk-free asset with a zero interest rate and multiple risky assets. Each asset i has a time-invariant supply of shares Q_i and pays a terminal dividend $D_{i,T} = D_{i,0} + \epsilon_{i,1} + ... + \epsilon_{i,T}$ at time T. The dividend innovations $\epsilon_{i,t}$ are observed by all traders at time t. Da, Huang, and Jin (2019) assume that the dividend innovations consist of a firm-specific and a market-wide component. We simplify their setting by assuming that all dividend innovations for all assets i come from a normal distribution with mean zero and variance σ^2. The dividend innovations therefore only have a firm-specific component and are identically and independently distributed over assets and time. If N is the number of risky assets in the economy, we can think of these risky assets as N time-series realizations of the same data-generating process. We drop the asset index i in the following for the sake of brevity.

Two types of agents with constant absolute risk aversion γ exist in the model: extrapolators and fundamental traders. The extrapolators' belief about the next price change of an asset in period t is given by:

$$\mathbb{E}_t^E[P_{t+1} - P_t] = \lambda_0 + \lambda_1 S_t. \tag{1}$$

The expected price change in the next period is a linear function of this period's sentiment, which is defined as

$$S_t = (1 - \lambda_2) \sum_{k=0}^{t-1} \lambda_2^k \left(P_{t-k} - P_{t-k-1}\right) + \lambda_2^t S_0. \tag{2}$$

return predictability tests with different combinations of asset classes, time periods, and definitions of extreme returns (see Amini, Gebka, Hudson, and Keasey, 2013, for a review). More recently, Savor (2012) provides a comprehensive large-sample investigation of short-term returns after extreme daily returns in the US. There is also an increased interest in examining extreme daily returns as a predictor for monthly returns (Bali, Cakici, and Whitelaw, 2011, Jiang and Zhu, 2017, Kumar, Ruenzi, and Ungeheuer, 2019).

Sentiment is a weighted average of past price changes and a starting value S_0. The parameter $0 < \lambda_2 < 1$ governs how strongly the extrapolator is influenced by recent price changes relative to more distant ones.

Fundamental traders form the second group of traders. They maximize their expected utility of wealth next period and assume that any mispricing will be corrected next period. The population of traders can be split into a fraction μ^E of extrapolators and a fraction $\mu^F = 1 - \mu^E$ of fundamental traders.

Defining $\alpha_t = \gamma\sigma^2 Q\left(T - t - 1 + \frac{1}{\mu^F}\right)$, Da, Huang, and Jin (2019) derive a closed-form solution for the equilibrium price of all assets in this economy. The price change of an asset is the difference between two adjoining prices and, in our setting, given by

$$
\begin{aligned}
&P_t - P_{t-1} = \\
&\frac{D_t + (\mu^F)^{-1}\mu^E[\lambda_0 + \lambda_1\left[(1-\lambda_2)\sum_{k=1}^{t-1}\lambda_2^k\left(P_{t-k} - P_{t-k-1}\right) + \lambda_2^t S_0\right] - \lambda_1(1-\lambda_2)P_{t-1}] - \alpha_t}{1 - (\mu^E/\mu^F)\lambda_1(1-\lambda_2)} \\
-&\frac{D_{t-1} + (\mu^F)^{-1}\mu^E[\lambda_0 + \lambda_1\left[(1-\lambda_2)\sum_{k=1}^{t-2}\lambda_2^k\left(P_{t-k-1} - P_{t-k-2}\right) + \lambda_2^{t-1} S_0\right] - \lambda_1(1-\lambda_{i,2})P_{t-2}] - \alpha_{t-1}}{1 - (\mu^E/\mu^F)\lambda_1(1-\lambda_2)} \\
&= \frac{1}{1 - (\mu^E/\mu^F)\lambda_1(1-\lambda_2)}\left[\epsilon_t + \gamma\sigma^2 Q\right] \\
&- \frac{(\mu^F)^{-1}\mu^E\lambda_1(1-\lambda_2)}{1 - (\mu^E/\mu^F)\lambda_1(1-\lambda_2)}\left[\sum_{k=1}^{t-1}(\lambda_2^{k-1} - \lambda_2^k)(P_{t-k} - P_{t-k-1})\right] \\
&- \frac{(\mu^F)^{-1}\mu^E\lambda_1}{1 - (\mu^E/\mu^F)\lambda_1(1-\lambda_2)}\left[(\lambda_2^{t-1} - \lambda_2^t)S_0\right]
\end{aligned}
\tag{3}
$$

The price change is a weighted sum of (i) the most recent dividend innovation ϵ_t plus the market risk premium $\gamma\sigma^2 Q$, (ii) the past price changes, and (iii) the starting sentiment S_0. The weights depend on the degree of return extrapolation (captured by the parameters λ_1 and λ_2) and the fraction of extrapolators μ^E in the economy.[2]

[2] Note that equation (3) holds for all assets in the economy. Differences in prices across assets are only caused by differences in dividend innovations.

Equation (3) suggests that price changes in period t are just a function of previous dividend innovations and the starting sentiment S_0. The following proposition states an exact formula for our setting. Appendix A contains the proof.

Proposition 1 (Price changes as a function of past epsilon shocks for $t \geq 2$)**.** *For* $S_0 = \gamma\sigma^2 Q$*,*

$$P_t - P_{t-1} = \left(\frac{1}{1-A}\right)\epsilon_t - \frac{A(1-\lambda_2)}{(1-A)^2}\sum_{j=1}^{t-1}\epsilon_j\left(\frac{\lambda_2 - A}{1-A}\right)^{t-1-j} + \gamma\sigma^2 Q \tag{4}$$

with $A = (\mu^E/\mu^F)\lambda_1(1-\lambda_2)$.

We assume $S_0 = \gamma\sigma^2 Q$ in proposition 1.[3] This assumption implies that extrapolators' belief about the mean of the first price change in the economy is a linear function of the market risk premium. For $\lambda_0 = 0$ and $\lambda_1 = 1$, the beliefs of all agents about the first price change coincide and are equal to the market risk premium.

An interesting special case is the parameterization $\lambda_2 = A$. Equation (4) in Proposition 1 becomes

$$P_t - P_{t-1} = \frac{1}{1-\lambda_2}\epsilon_t - \frac{\lambda_2}{1-\lambda_2}\epsilon_{t-1} + \gamma\sigma^2 Q. \tag{5}$$

We consider $\lambda_2 = A$ to be an attractive parameter choice for an application to daily data for two reasons. First, equation (5) implies that dividend innovations from the day before yesterday have no predictive power for price changes today, consistent with the idea that several days old information has no value for predicting price changes. Second, equation (5) further implies that this period's price change is normally distributed with mean $\gamma\sigma^2 Q - \frac{\lambda_2}{1-\lambda_2}\epsilon_{t-1}$ and variance $\frac{1}{(1-\lambda_2)^2}\sigma^2$. The mean of this period's price change is larger (smaller) than the market risk premium if the previous period's dividend innovation was negative (positive), consistent with the evidence on negative autocorrelations of daily returns. We therefore work with the assumption $\lambda_2 = A$, although our empirical implications are also valid for alternative parameterizations.[4]

[3] Equation (A.11) in Appendix A states the more involved formula for an arbitrary S_0.

[4] Appendix B provides an numerical example that allows for cross-sectional variation in the extrapolation parameters λ_1 and λ_2.

Equations (4) and (5) show that this period's dividend innovation ϵ_t has predictive power for the next price change $P_{t+1}-P_t$. A classic problem in financial economics is that researchers cannot directly observe dividend innovations. Even for salient information it is often hard to pin down the exact day the information hits the market (see, for example, the discussion in MacKinlay, 1997, among others), let alone the more difficult task to extract the exact piece of information that agents trade on each single day and for each asset.

However, in our setting, it is possible to use past prices to make inferences about the most recent dividend innovation. Solving equation (5) for ϵ_t gives

$$\epsilon_t = (1-\lambda_2)(P_t - P_{t-1}) + \lambda_2\epsilon_{t-1} - (1-\lambda_2)\gamma\sigma^2 Q. \tag{6}$$

It further holds for the dividend innovation in the previous period ϵ_{t-1},

$$\epsilon_{t-1} = (1-\lambda_2)(P_{t-1} - P_{t-2}) + \lambda_2\epsilon_{t-2} - (1-\lambda_2)\gamma\sigma^2 Q. \tag{7}$$

We can substitute the expression for ϵ_{t-1} from equation (7) into equation (6). Iterating this procedure for n periods, the formula for this period's dividend innovation becomes

$$\epsilon_t = \lambda_2^n\epsilon_{t-n} + (1-\lambda_2)\sum_{i=t-n+1}^{t}\lambda_2^{t-i}\left[(P_i - P_{i-1}) - \gamma\sigma^2 Q\right]. \tag{8}$$

An econometrician who observes an asset from the start of the economy is able to deduct all dividend innovations if the market risk premium ($\gamma\sigma^2 Q$) and the degree of return extrapolation (λ_2) is known and constant over time. However, time-constant parameters are unlikely, given the ample evidence on time varying risk aversion (Campbell and Cochrane, 1999), volatility jumps (Merton, 1976, Duffie, Pan, and Singleton, 2000), and time variation in investor sentiment (Baker and Wurgler, 2006).

A more realistic assumption is that the market risk premium and the degree of return extrapolation is approximately constant for a sufficiently short time period. Under this assumption, equation (8) suggests that past price changes from n previous periods can serve as a proxy for today's dividend innovation. Even without exact knowledge of γ, σ^2 and λ_2, assets that outperformed the market several times in a row ($(P_i - P_{i-1}) > \gamma\sigma^2 Q$ over the previous n periods) tend to have a higher dividend innovation ϵ_t in this period. Assets that have underperformed over the previous n periods tend to have a smaller dividend innovation ϵ_t. We call a situation where an asset has outperformed or underperformed the market *consecutively* over n previous periods a positive or negative streak of length n, respectively. We are now ready to formulate our main empirical predictions that we test in this paper.

Main Empirical Predictions: Assets which have experienced a negative (positive) streak in returns over the previous periods are expected to outperform (underperform) the market going one period forward. These abnormal returns tend to increase in absolute terms the longer the streak length.

3 Data

Theory provides little guidance on how long the empirical equivalent of one model period should be. As discussed in more detail in Da, Huang, and Jin (2019), the behavioral assumptions of the model imply quick reversals of sentiment and prices, and therefore seem more suitable for short-term data. We have argued in Section 2 that implications of the model are broadly consistent with existing empirical evidence collected on daily data. Our main analysis is therefore conducted with daily US data. The fact that our empirical implications can be tested using only past prices as data inputs allows us to examine international markets as well.

3.1 US Data

For the US, we use daily stock data from the Center for Research in Security Prices (CRSP). The data of stocks listed on the NYSE, AMEX or NASDAQ are collected over the period from January 1, 1998 to December 31, 2019.[5] We consider common stocks (share code 10 or 11) and exclude stocks quoted below $1 5 days prior to portfolio formation. Returns are calculated on the basis of closing prices and dividends. We adjust the return using the delisting return provided by CRSP. In order to test whether our results are driven by bid-ask bounces, we also calculate portfolio returns on the basis of mid-quotes using end-of-day bid and ask prices as well as dividends. Following Nagel (2012), we control for possible data inconsistencies by removing any bid-ask spread and midpoint return observation where the bid-to-midpoint ratio is smaller than 50% and the percentage point difference between midpoint return and closing price return is larger than 100% or smaller than -50%.

Streaks in returns can be identified by considering performance relative to the market over the past few trading days.[6] A streak in returns of length n occurs when the stock outperforms or underperforms the market consecutively between $t-1$ and $t-n$, where n is in the set {2, 3, 4, 5}.[7]

Quarterly earnings announcements are collected from Compustat and matched to the stocks in our sample. There are 399,511 successfully matched earnings announcements. The dates of the earnings announcements are shifted by 1 day if the trading volume is significantly higher on the following day, to account for after-hour announcements.

3.2 International Data

In an out-of-sample test, we confront our empirical implications with data from the regions Japan, Asia Pacific, Europe and Canada. The regions investigated are based on Fama and French (2012), with the exception that we investigate Canada separately from the US.

[5] We exclude data before 1998 to concentrate on a period with lower transaction costs and lower bid-ask spreads (Barclay, Christie, Harris, Kandel, and Schultz, 1999). This choice follows Nagel (2012), who argues that changes in order-handling rules on Nasdaq implemented in 1997 could influence the serial autocorrelation and lead to regime switches within the sample. However, there is a clear time-trend in the sample, indicating that transactions costs and bid-ask spreads decrease substantially over time. We will refer to this time trend repeatedly in the following sections.

[6] The market return as well as other factors are downloaded from Ken French's website.

[7] As an alternative specification, we estimate rolling betas based on the past 252 trading days and compare a stock's return to the ex-ante beta-forecast times the market return.

We collect daily market values and return data from Datastream from January 1, 1997 to October 29, 2017 for all stocks that have their primary listing in the regions Japan, Asia Pacific, Europe and Canada. In order to correct for data mistakes, we filter the data following Ince and Porter (2006) and Schmidt, Von Arx, Schrimpf, Wagner, and Ziegler (2017). Furthermore, we expand the word searches from Ince and Porter (2006) by the country specific screens reported in Griffin, Kelly, and Nardari (2010). We adjust the monthly dynamic filters implemented by Ince and Porter (2006) for daily data along the lines of Jacobs (2016) so that returns higher than 300% are set to "missing," as well as returns for which Ret_t and Ret_{t-1} are larger than 100%, and $(1 + Ret_t)(1 + Ret_{t-1}) - 1$ is smaller than 50% (Ince and Porter, 2006). Furthermore, companies that account for more than 90% of the entire market capitalization of the country are eliminated. Micro-cap stocks, i.e. stocks with an end-of-month unadjusted price below the 5% quantile of the domestic price distribution, are removed from the dataset. All observations are expressed in USD, and the market values and returns reported in other currencies are converted to USD using exchange rates from Datastream.[8]

Streaks in daily returns are identified in the same manner as those on the US market. The market-adjusted return is the raw return of the stock minus the market return of the country in which the stock is listed. The market return for each country is calculated as the value-weighted portfolio return of all stocks primarily listed in that country.

4 Empirical Results

In this section, we report our main tests of the empirical implications derived in Section 2. The model suggests that streaks in daily returns have predictive power for the return of the following day.

[8] During the considered time period, on January 2, 2002, many EU countries have switched to the Euro. In Datastream, all stocks delisted before the switch in 2002 still have their values reported in the original currency, whereas stocks that have delisting dates following the currency change or are still active are reported in Euro.

4.1 Streaks and Return Predictability

We form 10 portfolios based on the sign and length of previous n-day streaks and assume that stocks are bought at the formation day's closing price.[9] In what follows, we focus on value-weighted portfolios for various reasons. Value-weighted portfolios largely avoid biases caused by microstructure frictions (Blume and Stambaugh, 1983, Asparouhova, Bessembinder, and Kalcheva, 2013). Previous explanations of short-term autocorrelation like non-trading periods or non-synchronous trading (see, e.g., Boudoukh, Richardson, and Whitelaw, 1994) are less likely to apply for larger and more actively traded stocks. Classic short-term reversal returns tend to be substantially weaker for value-weighted portfolios, making value-weighted portfolios an ex-ante more challenging testing ground. Trading costs, particularly market impact costs, tend to be smaller for larger stocks.

Table 1, Panel A, reports the time-series average of value-weighted portfolio returns in excess of the market. Each portfolio is formed on a different streak length. After negative streaks in returns, portfolio returns increase with streak length; analogously, a longer positive streak is associated with lower returns.[10] For example, if we build a value-weighted portfolio with stocks that have underperformed the market on all the five previous trading days, the

[9] A difficulty is that we need to know the closing price at the end of a day in order to determine whether or not a streak in returns is intact. At the same time, though, a straight-forward testing procedure would assume that a stock can be bought at the closing price. This constitutes a timing problem, which can be dealt with in at least two ways. First, one can construct a portfolio shortly before the closing auction, including only those stocks that have extreme returns today that are unlikely to break streaks during the closing auction. From a theoretical point of view, such stocks are also those with the highest expected return in absolute terms over the holding period. Second, the strategy can be implemented by placing limit-on-close (LOC) orders for stocks whose streaks have a substantial probability to break in the closing auction. For example, let's assume that a stock has had a four-day negative streak in returns and has closed at \$100.01 on the previous day. Shortly before bids for the closing auction must be submitted, the stock is trading at \$99.99 and the market is expected to earn a zero return today. Assuming that the market return is indeed zero, the streak will be intact if the closing price is smaller than or equal to \$100. A LOC order with a buying limit of \$100 ensures that the stock will only be bought if the streak in daily returns continues. A recent academic paper that discusses order types in closing auctions at the US stock exchanges in detail is Comerton-Forde and Putniņš (2011).

[10] These portfolio returns are qualitatively similar to the price changes that we observe in simulated data using the model from Section 2, see Appendix B, Table B.4.

Table 1: Value-weighted market-adjusted returns of streak portfolios

For portfolios with different streak lengths, this table reports time-series averages of the value-weighted market-adjusted portfolio returns, the corresponding Newey-West t-statistics, and the number of stocks in the portfolio. The length of a streak is measured in number of days ranging from 1 to 5. Portfolio returns are based on closing prices in Panels A and on mid-quote prices in Panel B. No. of stocks reports the average number of firms in each portfolio per day. Portfolio returns are reported in percentages. Panel C reports excess-market returns of portfolios formed in the universe of small (below NYSE median market cap) and big stocks.

		Negative streaks					Positive streaks				
		Length of streak (days)					Length of streak (days)				
		1	2	3	4	5	1	2	3	4	5
Panel A: Value-weighted excess-market portfolio returns (in %)											
Avg. ex-mkt return		0.008	0.04	0.062	0.097	0.114	-0.008	-0.035	-0.05	-0.057	-0.077
t-stat		2.083	6.403	7.028	7.143	6.964	-2.11	-5.415	-5.54	-5.767	-5.851
Avg. no. of stocks		2288	1135	557	274	134	2147	995	460	214	101
Panel B: Value-weighted excess-market portfolio returns using midquotes (in %)											
Avg. ex-mkt return		-0.004	0.029	0.049	0.083	0.103	0.003	-0.024	-0.04	-0.046	-0.065
t-stat		-0.911	4.571	5.566	6.367	6.018	0.77	-3.726	-4.366	-4.732	-4.991
Avg. no. of stocks		2243	1112	546	268	132	2105	974	451	210	98
Panel C: Value-weighted excess-market portfolio returns among small (S) and big (B) stocks (in %)											
Avg. ex-mkt return	S	0.031	0.057	0.07	0.088	0.098	-0.01	-0.014	-0.012	-0.023	-0.035
t-stat		3.454	5.428	5.815	6.676	6.442	-1.006	-1.331	-1.161	-2.126	-2.824
No. of stocks		1783	883	433	212	104	1649	750	340	156	72
Avg. ex-mkt return	B	0.006	0.039	0.061	0.097	0.115	-0.007	-0.035	-0.052	-0.059	-0.082
t-stat		1.576	5.632	6.877	6.694	6.669	-1.896	-5.266	-5.491	-5.616	-5.753
No. of stocks		505	251	124	62	30	498	245	120	58	28

average market-adjusted return of this portfolio over the next trading day is 11.4 basis points, with a t-statistic of 6.964.[11]

In sum, these results strongly support the main empirical predictions of the theoretical model. The returns of value-weighted portfolios increase monotonically with streak length.

Table 2: Summary of long-short strategies

The long leg consists of the 4 value-weighted streak portfolios based on negative streaks stretching over 2 to 5 days, with each streak-length portfolio receiving 1/4 of the weight. The short leg consists of the 4 value-weighted portfolios based on positive streaks stretching over 2 to 5 days. Each column refers to a variant of a long-short streak strategy. They differ among four dimensions: (i) the prices used for return calculation (closing or mid prices), (ii) the threshold that separates daily winner and losers, (iii) if and how a portfolio hedges market risk exposure, and (iv) whether we use only the tercile of stocks with the lowest bid-ask spread for constructing the strategy. t-statistics are corrected for serial correlation and heteroskedasticity (Newey-West). The strategy beta is the coefficient of the full sample time-series regression of portfolio returns on the excess return of the market over the risk-free rate.

Return Calculation	close	mid	close	close	close	close	mid
Threshold	mkt	mkt	capm	mkt	capm	mkt	mkt
Hedge	unhedged	unhedged	unhedged	DT	DT	unhedged	unhedged
Subset						lowSpread	lowSpread
Avg. excess return (in %)	0.133	0.110	0.125	0.122	0.110	0.136	0.122
Std. dev. (in %)	1.091	1.104	0.925	0.890	0.877	1.175	1.171
t-stat	7.546	6.274	8.001	6.857	6.710	7.146	6.586
Annualized Sharpe ratio	1.935	1.578	2.151	2.177	1.986	1.841	1.651
Beta	0.218	0.224	0.143	0.098	0.018	0.217	0.218

The value-weighted portfolios based on a streak length of 2 to 5 days in Table 1 are used to create a dollar-neutral long-short portfolio. The short leg consists of the four portfolios with a positive streak of 2 to 5 days; the long leg consists of the portfolios with negative streaks of 2 to 5 days. In the long and the short leg, each streak-length portfolio receives the weight of $\frac{1}{4}$. Within each streak-length portfolio, stocks are still value-weighted. Descriptive statistics of the dollar neutral long-short portfolio are reported in Table 2. We calculate

[11] Table C.1 in Appendix C reports the CAPM, the Fama-French three-factor model (FF3), and the Fama-French five-factor model (FF5) alpha of the value-weighted streak portfolios. For example, the portfolio with stocks that have lost relative to the market during all previous 5 trading days yields an FF5-alpha of 12.2 basis points with a t-statistic of 6.713.

returns based on transaction prices and based on mid-quotes. The value-weighted portfolios earn 13.3 (t-stat: 7.546) and 11.0 (t-stat: 6.274) basis points on average, respectively. Based on transaction prices the strategy has an annualized Sharpe ratio of 1.935. Using mid-quotes reduces the annualized Sharpe ratio to 1.578.To put these numbers into perspective, Nagel's (2012) value-weighted industry reversal strategy calculated based on transactions prices earns 2 bps per day and has an annualized Sharpe ratio of 0.56.

We deliberately report the simplest version of streak strategies as our baseline specification. However, Table 2 contains further information on several variations. First, our baseline strategy classifies stocks into daily winners and losers based on a comparison of each stock's return with the market return. If the CAPM holds, it would be more sensible to compare each stock's return with the predicted return on that day according to the CAPM. In our baseline, we will systematically over-classify high-beta stocks as winners, as what looks like an outperformance could just be due to the fact that it is a high-beta stock and $1 \cdot Mkt$ is too low a threshold; and vice-versa for low-beta stocks. To implement the alternative CAPM threshold, we estimate for each stock and each day the CAPM betas by regressing stocks' daily excess returns on daily market excess returns. The estimation is based on the previous 252 days and we require at least 128 observations. The predicted excess return of a stock is then equal to the market excess return times the estimated beta. A positive streak in daily returns requires that the excess return of the stock is larger than the prediction of the CAPM. Table 2 shows that a strategy using a CAPM threshold is less volatile and has a higher annualized Sharpe ratio of 2.151.

Second, the long-short strategy in our baseline specification has a positive beta (0.218). It would be interesting to see how a portfolio performs that is closer to market neutrality. The CAPM-threshold version is one way to avoid systematically drawing market exposure by disproportionately often classifying high (low) beta stocks to the long (short) side. Indeed we see a drop in full-sample beta to 0.143 in column 3.

Alternatively, we can ex-ante hedge market exposure. We calculate returns of the hedged portfolio following the procedure of Daniel and Titman (1997, DT) and Daniel, Mota, Rottke, and Santos (2020b). In particular, for every day we calculate the preformation beta of the long-short portfolio. To do so, we hold the portfolio weights constant and calculate the hypothetical portfolio returns on each of the 252 trading days preceding the holding day. Subsequently, we regress these preformation returns on the excess return of the market to obtain the preformation beta. The preformation beta is used to calculate the returns of the hedged portfolio. Ex-ante hedging based on the above procedure decreases the systematic and total risk substantially and obtains an even higher annualized Sharpe ratio of 2.177.

4.2 Liquidity

One potential concern with a daily signal based on past short-term return is that it may proxy for illiquidity. We present several test results that are inconsistent with this hypothesis.

First, we weight stocks within portfolios by their previous day's market capitalization. Due to the skewed cross-sectional distribution of firm-size, the vast majority of weight is therefore put on large firms, which are less likely to be illiquid. Consistent with this, when we split the sample into small (S) and big (B) stocks at the NYSE median, and repeat the analysis within these subgroups, the results among big stocks (Table 1, Panel C) are almost exactly the same as those derived from the full sample. Interestingly, for longer streaks (and especially following positive streaks), results among small stocks are weaker than those among large stocks — inconsistent with the illiquidity intuition. Contrary to this, returns following the 1-day streak, which correspond to the classic short-term reversal signal, are stronger for small stocks, indicating that streaks and short-term reversal may be distinct phenomena. Furthermore, this result seems particularly surprising in light of recent papers emphasizing the role of small stocks, especially micro caps, in the anomaly literature (see, e.g., Hou, Xue, and Zhang, 2019).

Second, we re-calculate our results using midquote prices instead of closing prices (Table 1 Panel B). This tackles the worry that results may be influenced by bid-ask bounces that are driven by highly illiquid stocks, that rarely trade. This eliminates the significant market-adjusted return for the 1-day streak, i.e., daily reversals seem to be entirely driven by this phenomenon. For longer streaks, however, returns are smaller, but still highly statistically significant. Interestingly, it is especially the long side (i.e. buying stocks after negative streaks of 2 or more days) that is barely affected by using midquotes, with t-statistics between 4.6 and 6.4.

Third, we use the previous day's bid-ask half-spread to split the sample into three groups of illiquidity (Table 3, Panel A). Results in the low half-spread tercile remain strong and highly statistically significant for streaks longer than one day. Excess returns following 1-day streaks become insignificant for liquid stocks, while they are elevated for illiquid stocks (high bid-ask half-spread tercile). While returns following longer streaks are also elevated for illiquid stocks (3rd half-spread tercile, Panel A), that effect vanishes almost completely when using midquote prices instead (3rd half-spread tercile, Panel B). Liquid stocks' results are barely affected by the choice of returns calculated based on closing or midquote prices for any streak-length exceeding 1 day – again consistent with the idea that the streak phenomenon (streaks of at least 2 days) is distinct from short-term reversal (1-day past return signal).

Fourth, we regress the long-short streak portfolio on a host of "factors" (Table 4). When we include a daily rebalanced value-weighted portfolio going long illiquid stocks and short liquid stocks (based on the quoted bid-ask half-spread), it barely affects the alpha.

4.3 International Evidence

The theory-driven trading strategy based on streaks in returns exhibits significant daily abnormal returns when applied to U.S. stocks. In the following, we replicate these results on international stock markets, for two reasons. First, going outside the U.S. provides out-of-sample evidence (see, e.g., the discussion in Asness, Moskowitz, and Pedersen, 2013).

Table 3: Sorts on half-spread within each streak portfolio

Each day, the cross section of stocks is ranked based on their previous day's bid-ask half-spread and assigned into terciles. In each bucket, the streak portfolios are constructed for a streak length of 1 to 5 days and the resulting portfolios are value-weighted. The returns are market-adjusted and the t-statistics are Newey-West adjusted. No. of stocks reports the average number of stocks in the portfolio per day (there are no days where any portfolios is empty for this double-sort). Closing prices are used to calculate returns in Panel A, and mid-prices in Panel B.

Panel A: Closing-Prices

		After negative streaks					After positive streaks				
Half-spread		Length of streak (days)					Length of streak (days)				
Tercile		1	2	3	4	5	1	2	3	4	5
3 (high)	Ex. Mkt Ret. (in %)	0.088	0.133	0.162	0.207	0.245	-0.136	-0.163	-0.165	-0.172	-0.153
	t-stat	7.196	9.116	8.550	9.286	9.019	-10.262	-10.664	-9.577	-7.703	-5.419
	No. of stocks	767	371	176	83	40	693	296	125	53	23
2	Ex. Mkt Ret. (in %)	0.010	0.040	0.050	0.087	0.112	-0.006	-0.028	-0.036	-0.050	-0.045
	t-stat	1.421	4.266	4.224	5.462	6.050	-0.777	-3.140	-3.100	-3.648	-2.640
	No. of stocks	751	378	189	95	47	707	333	157	74	35
1 (low)	Ex. Mkt Ret. (in %)	0.005	0.038	0.061	0.102	0.118	-0.008	-0.036	-0.052	-0.059	-0.080
	t-stat	1.091	5.333	6.564	7.012	6.523	-1.595	-4.584	-4.768	-5.109	-5.542
	No. of stocks	740	371	185	92	46	719	352	172	84	41

Panel B: Mid-Prices

3 (high)	Ex. Mkt Ret. (in %)	-0.058	-0.034	-0.007	0.034	0.065	0.026	0.018	0.020	0.010	0.028
	t-stat	-4.717	-2.414	-0.409	1.683	2.649	2.195	1.320	1.264	0.490	1.084
	No. of stocks	767	371	176	83	40	693	296	125	53	23
2	Ex. Mkt Ret. (in %)	-0.009	0.019	0.030	0.061	0.085	0.019	-0.006	-0.014	-0.025	-0.023
	t-stat	-1.247	1.998	2.537	4.126	4.689	2.491	-0.679	-1.169	-1.852	-1.361
	No. of stocks	751	378	189	95	47	707	333	157	74	35
1 (low)	Ex. Mkt Ret. (in %)	-0.002	0.031	0.053	0.093	0.110	-0.003	-0.030	-0.046	-0.052	-0.071
	t-stat	-0.335	4.582	5.728	6.725	5.995	-0.591	-3.936	-4.321	-4.692	-5.054
	No. of stocks	740	371	185	92	46	719	352	172	84	41

Table 4: Alpha of streak portfolios controlled for the Fama-French five-factors and the short-term reversal factor

This table reports results of regressions of the value-weighted streak portfolio long-short returns on various combinations of the five Fama-French factors, the short-term reversal factor, and the IVOL factor. Newey-West t-statistics are reported in parentheses next to the estimates. The factors are taken from Kenneth French's website: MktRF is the market portfolio minus the risk free rate; SMB is the small-minus-big factor; HML is the high-minus-low book-to-market factor; RMW is the robust minus weak profitability factor, and CMA is the conservative-minus-aggressive investment factor. More details on the factors can be found in Fama and French (1993, 2015). STR is the short-term reversal factor. To construct the IVOL factor, we sort all stocks based on the idiosyncratic volatility of the previous 21 (at least 12) trading days each day. IVOL is the return of a value-weighted portfolio long the quintile of stocks with the highest and short the quintile of stocks with the lowest idiosyncratic volatility (Ang, Hodrick, Xing, and Zhang, 2006). BASPR is a portfolio long the quintile of stocks with the highest and short the quintile with the lowest bid-ask-spread. WMOM is the value-weighted weekly momentum factor from Gutierrez and Kelley (2008). The intercept is reported in percent.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	0.127 (6.319)	0.137 (6.728)	0.105 (6.582)	0.112 (6.400)	0.112 (6.928)	0.111 (7.030)	0.073 (6.380)
MktRF	0.218 (7.160)	0.130 (6.243)		0.032 (1.308)	0.042 (1.718)	0.030 (1.030)	0.086 (4.232)
HML	-0.127 (-1.824)	0.028 (0.527)		0.029 (0.552)	0.035 (0.639)	0.045 (0.820)	-0.025 (-0.555)
SMB	-0.016 (-0.432)	-0.053 (-1.515)		-0.036 (-1.029)	-0.013 (-0.326)	0.002 (0.043)	-0.018 (-0.683)
RMW		-0.224 (-3.889)		-0.140 (-2.588)	-0.197 (-3.325)	-0.200 (-3.366)	
CMA		-0.424 (-3.958)		-0.260 (-2.977)	-0.302 (-3.663)	-0.294 (-3.554)	
STRev			0.563 (10.980)	0.495 (11.147)	0.500 (11.617)	0.503 (11.746)	
IVOL					-0.038 (-1.302)	-0.039 (-1.369)	
BASPR						-0.037 (-1.000)	
WMOM							-0.460 (-30.596)
R^2	0.0633	0.0944	0.2146	0.2341	0.2352	0.2359	0.5257

Second, a larger stock universe would allow us to develop a portfolio that achieves a higher level of diversification.

We analyze four additional international regions: Asia Pacific, Canada, Europe, and Japan, which we have chosen following Fama and French (2012). Unlike in their work, we do not include the U.S. stock market in the same region as Canada, because we have already given a detailed summary of the streak strategy in the U.S. stock market. For international markets, we use data from 1998-2019, as in the U.S. analysis.

In all regions, similar return predictability after streaks in daily returns can be observed. Panel A of Table 5 reports the value-weighted market-adjusted returns after 1- to 5-day streaks. In all regions, we see sizeable returns. Absolute portfolio returns on the holding day increase with streak length in the predicted direction. In terms of magnitude, absolute returns in international markets tend to be slightly smaller than in the U.S. This result is broadly consistent with Jacobs (2016) and Jacobs and Müller (2020), who report that stock market anomalies are often of comparable magnitude and sometimes even less pronounced in international markets.

The U.S. and international streak portfolios do not exhibit highly correlated returns, as reported in Panel B of Table 5. This suggests further benefits from diversifying internationally. Indeed, a diversified streak strategy that puts the same weight on the regions Japan, U.S., Canada, Europe and Pacific Asia increases the annualized Sharpe ratio to 3.25. A strategy based solely on U.S. stocks has an annualized Sharpe ratio of 1.94 in the same time period.

5 Further Empirical Evidence

In this section, we gather further empirical evidence on the U.S. streak strategy. First, we show that streak strategies are different from well-known short-term strategies. Second, we estimate the model parameter λ_2 empirically on a rolling basis and show that stocks which

Table 5: International streak portfolios

Panel A presents the value-weighted streak portfolio returns for different streak lengths for 4 regions: namely, Asia Pacific, Canada, Europe, and Japan. For each streak portfolio with different streak lengths of positive and negative returns, this table reports time-series averages of value-weighted market-adjusted returns, Newey-West t-statistics, and no. of stocks in the portfolio. The length of a streak is measured in number of days, ranging from 1 to 5. The values reported are those recorded on the day following the streak. No. of stocks is the average number of stocks in each portfolio on each day. Portfolio returns are reported in percentage.

Panel B presents the correlation matrix of the regional long-short streak portfolios, including those of the US. In each region, the long-leg consists of the four portfolios with streaks of 2 to 5 days in negative returns; the short-leg consists of the portfolios with streaks of 2 to 5 days in positive returns. The last row presents the annualized Sharpe ratio of each regional strategy as well as that of a combination strategy, that invests $\frac{1}{5}$ into each region.

Panel A: Value-weighted market-adjusted portfolio returns (in %)

	Negative streaks Length of streak (days)					Positive streaks Length of streak (days)				
	1	2	3	4	5	1	2	3	4	5
Region: Asia Pacific										
Avg. ex-mkt return	-0.008	0.046	0.080	0.110	0.153	0.006	-0.028	-0.062	-0.081	-0.094
t-stat	-1.246	5.658	8.484	8.822	8.848	0.932	-3.401	-5.843	-5.499	-4.715
No. of stocks	1153	568	274	130	62	1023	442	190	82	36
VW quoted half-spread	0.325	0.343	0.370	0.418	0.488	0.319	0.331	0.352	0.390	0.467
Region: Canada										
Avg. ex-mkt return	0.003	0.049	0.083	0.117	0.167	0.011	-0.009	-0.042	-0.071	-0.135
t-stat	0.271	3.185	4.433	5.523	5.370	0.838	-0.621	-2.451	-3.233	-3.869
No. of stocks	391	188	89	41	20	364	161	71	31	14
VW quoted half-spread	0.239	0.271	0.320	0.394	0.508	0.228	0.247	0.286	0.351	0.455
Region: Europe										
Avg. ex-mkt return	0.004	0.041	0.065	0.094	0.120	0.000	-0.023	-0.050	-0.068	-0.082
t-stat	0.744	6.529	7.813	8.842	9.276	0.039	-2.909	-5.056	-6.750	-6.794
No. of stocks	2016	997	488	238	116	1848	833	374	169	77
VW quoted half-spread	0.198	0.200	0.205	0.218	0.243	0.193	0.190	0.189	0.196	0.211
Region: Japan										
Avg. ex-mkt return	-0.008	0.032	0.076	0.110	0.136	0.013	-0.015	-0.040	-0.064	-0.085
t-stat	-1.376	4.551	8.393	9.675	9.063	2.103	-2.058	-3.989	-4.972	-5.480
No. of stocks	1233	616	303	147	71	1144	527	240	108	49
VW quoted half-spread	0.157	0.164	0.169	0.175	0.181	0.157	0.163	0.169	0.174	0.181

Panel B: Correlation and Sharpe ratios of international streak strategies

	Asia Pacific	Canada	Europe	Japan	U.S.	Combination
Asia Pacific	1	0.1	0.17	0.16	0.09	
Canada	0.1	1	0.2	0.05	0.32	
Europe	0.17	0.2	1	0.11	0.33	
Japan	0.16	0.05	0.11	1	0.05	
U.S.	0.09	0.32	0.33	0.05	1	
Sharpe Ratio	2.10	1.63	2.09	1.85	1.93	3.25

are subject to a higher degree of return extrapolation earn higher streak returns. Third, we test whether the streak strategy remains profitably after accounting for trading costs. This is especially important for a trading strategy that is rebalanced daily. Fourth, we look at streak strategy returns on earnings announcement days.

5.1 Distinctness from known short-term reversal strategies

Figure 1 contains four panels that show the development of streak returns and several short-term reversal strategies over time. These panels highlight differences and similarities among them.

Panel A reports the six-month moving average of the value-weighted long-short portfolio returns as constructed in Section 4. Panel A also contains the six-month moving average of the value-weighted quoted half-spread on the holding day. We follow the market microstructure literature and compute the quoted half-spread of a stock as the difference between the quoted bid and the quoted ask price divided by two times the mid price (see, e.g., Bessembinder and Venkataraman, 2010). If trades are executed at quoted ask and bid prices, quoted half-spreads can be viewed as a proxy for the one-way transaction costs of a trade. CRSP quoted spreads closely approximate TAQ effective spreads, especially in recent years (Chung and Zhang, 2014, Abdi and Ranaldo, 2017). On each day, the value-weighted half-spread of the long-short portfolio is computed as a simple average of the ten value-weighted quoted half-spreads of all portfolios that were combined to the long-short streak portfolio. Panel A shows a six-month moving average of this time-series.

Half-spreads are high in the late 90's. In the early 2000's, spreads significantly decline as a result of the decimalization of quotes (Bessembinder, 2003) and the rise of algorithmic trading (Hendershott, Jones, and Menkveld, 2011). Afterwards, they continue to decrease at low levels with a temporary rise after the market decline during the financial crisis (Hameed, Kang, and Viswanathan, 2010). The overall pattern may be explained with a limits of arbitrage story. Arbitraging away mispricing caused by extrapolation was too costly in the

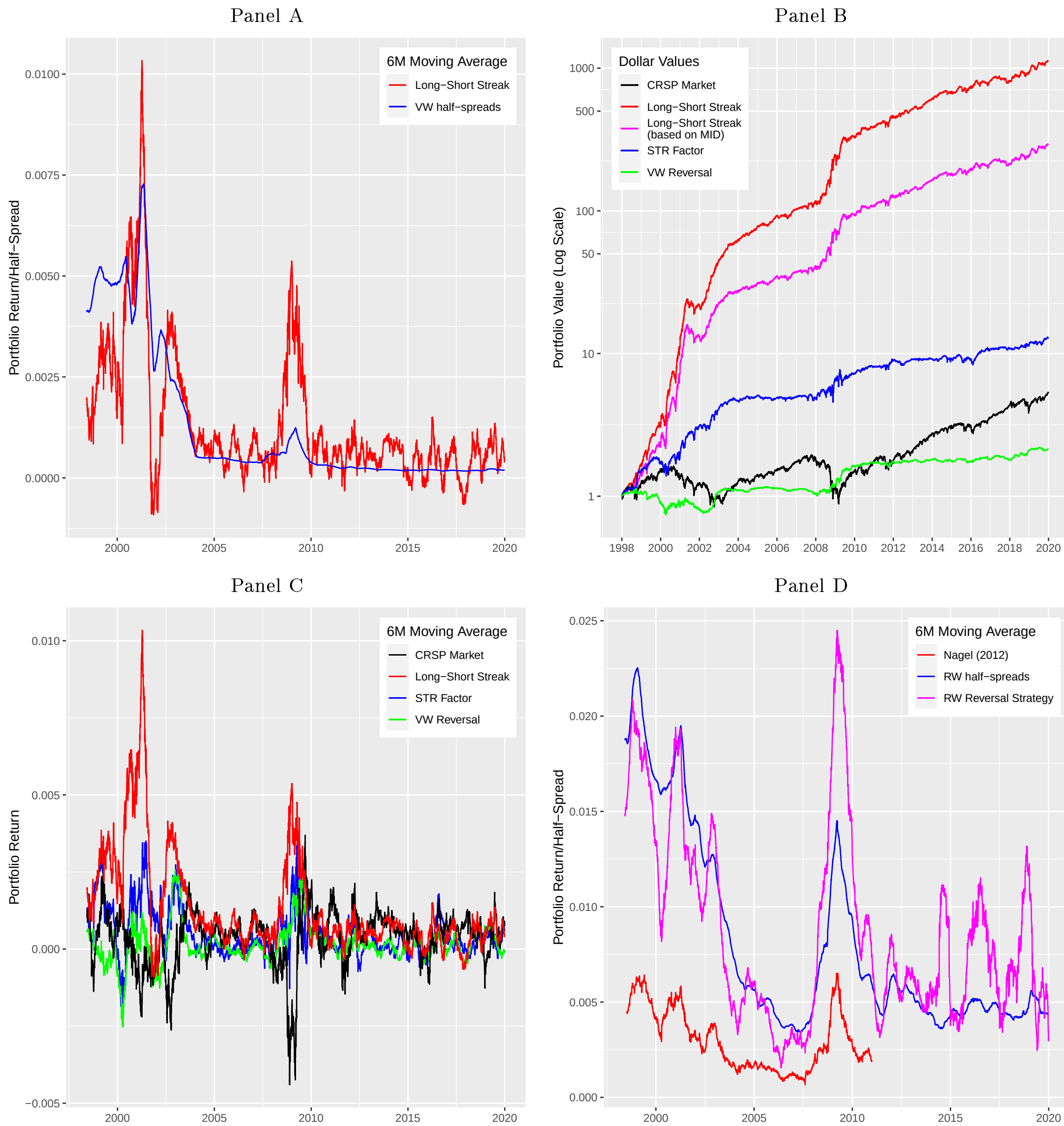

Figure 1: Returns and spreads over time

Panel A reports the six-month moving average of daily returns and of the average value-weighted quoted half-spreads of the long-short streak portfolio. The VW half-spread on each day is computed as a simple average of the ten value-weighted quoted half-spreads of all portfolios that were combined to long-short streak strategies. The weighting schemes of returns and spreads in Panel A are identical. Panel B plots how a dollar invested in four strategies at the beginning of 1998 would have developed over time before costs. The long-short streak strategy is plotted based on end-of-day returns. The graph also shows the development for the value-weighted market, the short-term reversal factor (STR), and a value-weighted portfolio that goes long (short) all stocks that have underperformed (outperformed) the market over the previous day (VW Reversal). Panel C reports the six-month moving average of daily streak returns (as in Panel A), but adds the market portfolio, the STR factor, and the VW Reversal. Panel D shows the six-month moving averages of Nagel's (2012) reversal strategy (Data source: https://voices.uchicago.edu/stefannagel/code-and-data/, last accessed: September 2019) and of the return-weighted (RW) average of quoted half-spreads. The RW reversal strategy forms a portfolio that goes long (short) the stocks that have underperformed (outperformed) the market over the previous day, and weights each stock with the absolute value of the market-adjusted return from the previous day.

late 90's. After spreads have fallen in the 2000's, streak returns have decreased as well. The pattern in Panel A is consistent with Chordia, Subrahmanyam, and Tong (2014), who report that anomalies are in general attenuated after decimalization. They argue that this attenuation is caused by decreased trading costs.

Panel B of Figure 1 plots the development of $1 invested in the long-short streak portfolio. The figure further compares our long-short streak strategy to the market portfolio, a value-weighted long-short portfolio that goes long all previous day's losers and short all previous day's winners (VW Reversal), and the short-term reversal (STR) strategy from Kenneth French's data library. Panel C compares these strategies in terms of their six-month moving average returns.

On a value-weighted basis, our streak strategy outperforms the well-known Fama-French short-term reversal factor by a wide margin. This is a fair comparison, as the STR factor, like our long-short portfolio, is a combination of value-weighted portfolios. Note that the outperformance of the streak portfolio is not achieved by overweighting small stocks; to the contrary, our long-short portfolio invests in larger firms than the STR factor on average (based on our replication of the STR factor; not reported).

Panel D shows the six-month moving average return of a return-weighted reversal strategy. This strategy overweights small stocks and is therefore conceptually different from the strategies that partly or completely follow value-weighted approaches. Not surprisingly, returns and spreads for these smaller stocks are much higher. We also plot Nagel's (2012) reversal strategy, which is an equally-weighted combination of ten return-weighted reversal portfolios. Return-weighted strategies are highly correlated with return-weighted quoted half-spreads over the entire sample. This observation is consistent with the idea that return-weighted strategies solely capture returns from liquidity provisions (as argued by Nagel, 2012), while streak strategies do not. We further see that a simple value-weighted reversal strategy underperforms the market (Panel B). Value-weighted strategies tend to perform well when the market does poorly (Panels B and C).

We want to make sure that our strategy is conceptually and economically different from established short-term reversal strategies, beyond the qualitative discussion of the patterns visualized in Figure 1. Therefore, we conduct four empirical tests.

First, our long-short streak portfolio generates a daily alpha of 11.2 basis points after controlling for the five Fama-French factors and the short-term reversal factor (see Table 4). If our strategy is just a different way of calculating the well-known reversal-factor portfolio, we would observe an alpha of zero. The strategy also does not load significantly on the IVOL-factor (Ang, Hodrick, Xing, and Zhang, 2006), and the alpha consequently remains unchanged. The loading on weekly momentum (Gutierrez and Kelley, 2008) is strongly and significantly negative (column 7). The alpha goes down to 7.3 basis points, but the t-statistic of the alpha remains large (6.380).

Second, we investigate if our streak strategy solely captures that stocks with extreme $t-1$ returns have more extreme reversal returns on the following day. We match streak stocks to stocks without streaks on size and return in $t-1$. As in Barber and Lyon (1997), we first find all stocks with a market value between 70% and 130% of the market value of the streak stock, choosing then from this group of possible control stocks those with the closest returns on day $t-1$ compared with the streak stocks. If the streak strategy is only capturing the reversal of a stock with very high return in $t-1$, then portfolios of the streak stocks and their matched stocks would not have significantly different returns.

In Table 6, we report the average daily value-weighted portfolio return on the holding day for streak stocks and control stocks. We report the value-weighted returns after negative and positive streaks in returns. Streak length is varied from 2 to 5 days. For example, if a stock has a 5-day negative streak in returns, then the market-adjusted returns on all 5 days prior to t are negative; however, the matched control stock has a negative market-adjusted return only on day $t-1$, but a positive one on day $t-2$. We report the difference in returns between the streak stock portfolio and control stock portfolio in the last column of Table 6.

Table 6: Excess-market returns of streak stocks vs. control stocks

This table reports the holding day returns of stocks with streaks in returns and those of matched control stocks. Matching controls for size and the most recent daily return using the following two step procedure. First, the universe of possible control stocks is defined as all firms where (i) the excess-market return on the day before yesterday has the opposite sign of yesterday's market-adjusted return and (ii) where the market capitalization lies between 70% and 130% of the market capitalization of the streak stock. Second, stocks with streaks in returns are matched to potential control stocks based on the most recent daily return. The table reports the average value-weighted excess-market return for a portfolio of streak stocks and a portfolio of matched control stocks. The length of a streak is measured in number of days ranging from 2 to 5. The sign of the streak return is reported in the second column. t-statistics are reported in parentheses next to average returns and are Newey-West-adjusted.

	After negative streaks			
	2	3	4	5
Treatment	0.040 (6.403)	0.062 (7.027)	0.097 (7.143)	0.114 (6.965)
Control	-0.022 (-3.084)	-0.020 (-2.559)	-0.017 (-1.951)	-0.013 (-1.364)
Diff	0.062 (5.785)	0.082 (6.172)	0.114 (6.587)	0.127 (6.384)
	After positive streaks			
	2	3	4	5
Treatment	-0.035 (-5.415)	-0.050 (-5.542)	-0.057 (-5.766)	-0.077 (-5.880)
Control	0.018 (2.572)	0.027 (3.339)	0.020 (2.154)	0.031 (2.721)
Diff	-0.053 (-4.874)	-0.076 (-5.912)	-0.077 (-5.725)	-0.108 (-5.877)

On holding day t, returns of streak portfolios are consistently more extreme than returns of control portfolios. All differences are highly statistically significant. We further observe that, for both negative and positive streaks, differences in the absolute value-weighted portfolio returns increase with streak length.

Our third test deals with the possibility that streak strategies are just picking up large liquidity shocks among highly illiquid stocks. The premise would then be that those shocks need more than a day to find their way into prices. We consider this possibility to be ex-ante unlikely, given the excess returns that we have already reported for value-weighted portfolios. In a value-weighted portfolio, the stocks with the highest weights tend not to be illiquid.

Table 7: Streaks and cumulative return Fama-MacBeth regressions

This table reports the Fama-MacBeth regression coefficients for eight different regressions. In each regression, the market-adjusted return at time t is the dependent variable. A positive streak dummy, a negative streak dummy, and the cumulative excess return for the n-days before t are the independent variables. In Panels B and C, the excess return on the previous day (Ret-Rf_{t-1}), or the bid-ask half-spread on the previous day, are added as an additional control variable, respectively. The regressions are conducted for variables based on an n of 2 to 5 days. All Panels report time-series averages of coefficients of cross-sectional weighted least-square regressions, where the market cap on the previous day is used as the weight. t-statistics based on Newey-West-adjusted standard errors are reported in parentheses.

	2-days	3-days	4-days	5-days
Panel A				
Intercept	-0.718 (-2.831)	-0.862 (-3.939)	-0.870 (-4.651)	-0.904 (-5.281)
n-day positive streak	-0.013 (-2.348)	-0.023 (-3.086)	-0.020 (-2.406)	-0.032 (-3.328)
n-day negative streak	0.025 (4.255)	0.039 (5.719)	0.066 (6.309)	0.067 (5.401)
n-day cumulative excess return	-0.368 (-2.905)	-0.441 (-4.032)	-0.446 (-4.759)	-0.462 (-5.423)
Panel B				
Intercept	-1.515 (-3.974)	-1.298 (-5.140)	-1.176 (-5.869)	-1.111 (-5.993)
n-day positive streak	-0.011 (-1.977)	-0.020 (-2.863)	-0.017 (-2.125)	-0.028 (-2.904)
n-day negative streak	0.024 (4.417)	0.038 (5.792)	0.063 (6.290)	0.065 (5.384)
n-day cumulative excess return	-0.767 (-4.033)	-0.660 (-5.209)	-0.599 (-5.985)	-0.567 (-6.126)
$Ret - Rf_{t-1}$	0.601 (2.251)	0.381 (1.821)	0.260 (1.334)	0.143 (0.746)
Panel C				
Intercept	-0.539 (-2.088)	-0.702 (-3.216)	-0.746 (-4.033)	-0.798 (-4.663)
n-day positive streak	-0.013 (-2.287)	-0.025 (-3.286)	-0.022 (-2.718)	-0.036 (-3.689)
n-day negative streak	0.027 (4.682)	0.041 (5.928)	0.070 (6.388)	0.073 (5.733)
n-day cumulative excess return	-0.280 (-2.176)	-0.362 (-3.312)	-0.384 (-4.153)	-0.410 (-4.821)
Bid-ask halfspread	-0.569 (-0.671)	-0.706 (-0.863)	-0.728 (-0.869)	-0.703 (-0.836)

Table 8: Sorts on t-1 returns within each streak portfolio

Each day, the cross section of stocks is ranked based on their previous day's return and assigned into terciles. In each bucket, the streak portfolios are constructed for a streak length of 1 to 5 days and the resulting portfolios are value-weighted. The returns are market-adjusted and the t-statistics are Newey-West adjusted. No. of stocks reports the average number of stocks in the portfolio for all days for which the portfolio is not empty. No. of days refers to the number of days for which the portfolio is not empty.

		After negative streaks					After positive streaks				
RET_{t-1}		Length of streak (days)					Length of streak (days)				
Tercile		1	2	3	4	5	1	2	3	4	5
3 (high)	Ex. Mkt Ret. (in %)						-0.010	-0.038	-0.058	-0.074	-0.100
	t-stat						-1.687	-4.699	-5.424	-6.465	-6.685
	No. of stocks						1493	682	315	147	69
	No. of days						5535	5535	5535	5535	5535
2	Ex. Mkt Ret. (in %)	-0.006	0.014	0.022	0.050	0.041	-0.004	-0.024	-0.034	-0.031	-0.037
	t-stat	-1.333	2.163	2.336	3.967	2.750	-0.839	-3.409	-3.524	-2.572	-2.665
	No. of stocks	824	422	207	100	49	675	322	150	70	33
	No. of days	5455	5455	5453	5450	5441	5356	5356	5352	5345	5324
1 (low)	Ex. Mkt Ret. (in %)	0.019	0.059	0.090	0.132	0.170					
	t-stat	2.855	6.334	7.865	7.87	8.029					
	No. of stocks	1469	716	352	174	86					
	No. of days	5535	5535	5535	5535	5535					

To address the argument formally, we run market-value-weighted Fama-MacBeth regressions. Streak dummies always remain significant, no matter if we include just the market-adjusted return of the previous day or additionally the previous n-day market-adjusted returns (see Table 7).

Our fourth and final test first sorts the cross-section of stocks into three terciles based on the $t-1$ market-adjusted return each day. Afterwards, we calculate streak strategies within these terciles. Table 8 reports the value-weighted market-adjusted returns on the holding day. We find that value-weighted portfolios in the upper and lower $t-1$ return tercile still have significant holding day returns after streaks. This provides further evidence that streaks in returns do not simply capture the predictability of stocks with high absolute $t-1$ returns.

5.2 Estimating the level of extrapolation λ_2

The model suggests that the holding day returns after streaks in daily returns should be higher for stocks with a higher λ_2 parameter. Recall that the parameter λ_2 determines the weight extrapolators put on more recent price changes compared to price changes that happened further in the past. To test whether stocks with a higher λ_2 have a higher holding day return after streaks, we estimate an empirical λ_2 for each stock on each sample day.

We go back to the model and rearrange equation (8) to write $P_t - P_{t-1}$ as a function of the last 5 price changes, so that n = 6. This leads to the following specification of $P_t - P_{t-1}$:

$$\begin{aligned} P_t - P_{t-1} = Q\gamma\sigma^2 &- \lambda_2^5\left(P_{t-5} - P_{t-6} - Q\gamma\sigma^2\right) - \lambda_2^4\left(P_{t-4} - P_{t-5} - Q\gamma\sigma^2\right) \\ &- \lambda_2^3\left(P_{t-3} - P_{t-4} - Q\gamma\sigma^2\right) - \lambda_2^2\left(P_{t-2} - P_{t-3} - Q\gamma\sigma^2\right) \\ &- \lambda_2\left(P_{t-1} - P_{t-2} - Q\gamma\sigma^2\right) + \frac{\epsilon_t - \lambda_2^6\epsilon_{t-6}}{1-\lambda_2}. \end{aligned} \tag{9}$$

A time-series regression is run separately for each stock i on each sample day s using the previous 252 trading days (indexed by t) as observations:

$$\begin{aligned} AdjRet_{i,t} = \alpha_i &+ \beta_{i,s,1}AdjRet_{i,t-1} + \beta_{i,s,2}AdjRet_{i,t-2} + \beta_{i,s,3}AdjRet_{i,t-3} \\ &+ \beta_{i,s,4}AdjRet_{i,t-4} + \beta_{i,s,5}AdjRet_{i,t-5} + u_i \end{aligned} \tag{10}$$

where $AdjRet_{i,t}$ is the market-adjusted return of stock i on day t. Equations (9) and (10) allow to set-up the following minimization problem to estimate $\hat{\lambda}_{i,s,2}$ of stock i given the estimated $\hat{\beta_{i,s}}$'s from equation (10):[12]

$$\hat{\lambda}_{i,s,2} = \arg\min_{\lambda_{i,s,2}} \sum_{j=1}^{5} (\hat{\beta}_{i,s,j} + \lambda_{i,s,2}^{j})^2 \tag{11}$$

Using the estimated $\hat{\lambda}_{i,s,2}$ values, we separate stocks into terciles on each sample day s. We compute the streak strategy returns within each tercile. Table 9 reports the value-weighted returns. Consistent with the theory, we observe that streak strategy returns tend to be higher for stocks with a high $\hat{\lambda}_{i,s,2}$. For shorter streaks, strategies using only stocks in the highest $\hat{\lambda}_2$ tercile yield 5.6 to 6.6 basis points higher returns. This result allows to increase returns from streak strategies, using just the past prices from the previous trading year as data input.

5.3 Streak Strategy & Trading Costs

A growing number of recent studies investigates the trading costs faced by institutional investors (see, e.g., Keim and Madhavan, 1997, Korajczyk and Sadka, 2004, Lesmond, Schill, and Zhou, 2004, Frazzini, Israel, and Moskowitz, 2012, Engle, Ferstenberg, and Russell, 2012, De Groot, Huij, and Zhou, 2012, Novy-Marx and Velikov, 2015, Frazzini, Israel, and Moskowitz, 2018). Implementing a streak portfolio undoubtedly entails a large amount of daily trading.

[12] Note that the optimization problem is not necessarily globally convex. However, the objective function is locally concave for unusual parameter combinations only, like a high absolute value of $\hat{\beta}_{i,s,2}$. To minimize the probability of finding a local minimum and not the global one, we adopt a variant of graduated optimization, a method to solve non-convex optimization problems faced in machine learning and image processing (see Blake and Zisserman, 1987). We start by minimizing the simple objective function $(\hat{\beta}_{i,s,1} + \lambda_{i,s,2})^2$ and obtain the global minimum at $-\hat{\beta}_{i,s,1}$ analytically. In the next step, we solve the slightly more complicated optimization problem $\arg\min_{\lambda_{i,s,2}} (\hat{\beta}_{i,s,1} + \lambda_{i,s,2})^2 + (\hat{\beta}_{i,s,2} + \lambda_{i,s,2}^2)^2$ numerically, using $-\hat{\beta}_{i,s,1}$ as starting value. In the third iteration, we add the term $(\hat{\beta}_{i,s,3} + \lambda_{i,s,2}^3)^2$ to the objective function and use again the minimum value from the previous step as starting value. This step-wise procedure continues until we have minimized the objective function in equation (11).

Table 9: Relationship between streak strategy returns and estimated $\hat{\lambda}_2$

For the cross-section of US equities, rolling $\hat{\lambda}_2$ estimates (based on the past 252 trading days) are used to sort stocks into terciles. We build portfolios from the intersection with the five positive and negative streak portfolios. The table reports value-weighted portfolio returns and Newey-West t-statistics.

λ_2	After negative streaks				
Tercile	1	2	3	4	5
1 (Low)	0.025 (1.706)	0.060 (3.640)	0.092 (4.974)	0.122 (5.547)	0.140 (5.979)
2	0.045 (3.192)	0.076 (5.029)	0.089 (5.038)	0.122 (6.057)	0.128 (5.288)
3 (High)	0.088 (6.016)	0.124 (7.984)	0.125 (7.075)	0.166 (7.877)	0.198 (8.136)
High-Low	0.063 (8.232)	0.065 (6.360)	0.033 (2.495)	0.044 (2.387)	0.058 (2.705)
λ_2	After positive streaks				
Tercile	1	2	3	4	5
1 (Low)	0.036 (2.378)	0.005 (0.349)	-0.019 (-1.063)	-0.037 (-1.998)	-0.061 (-3.058)
2	0.021 (1.409)	-0.003 (-0.221)	-0.015 (-0.885)	-0.019 (-1.046)	-0.022 (-1.084)
3 (High)	-0.030 (-1.865)	-0.051 (-2.952)	-0.048 (-2.673)	-0.051 (-2.610)	-0.050 (-2.054)
High-Low	-0.066 (-8.522)	-0.056 (-6.102)	-0.029 (-2.486)	-0.014 (-0.899)	0.011 (0.503)

A full analysis of the question if daily return predictability could be economically exploitable is beyond the scope of the paper. However, we aim to get a rough idea on how close to profitable the strategy might be after accounting for transaction costs. We start approaching this question by recognizing that quoted half-spreads have dramatically decreased after decimalization. Assuming that trades occur at quoted bids and asks, quoted spreads equal the trading costs of a round-trip and half-spreads equal the costs of a single trade. We look at an equally-weighted average of the two value-weighted long portfolios consisting of stocks that have experienced a negative streak in daily returns of length 4 and 5, respectively. Consistent with the theory, Table 1 shows that these portfolios are the two best performing ones over the entire sample. Furthermore, long portfolios avoid short-selling costs.

To put the after-costs performance into perspective, we perform the following calculation:

$$r_t^{AC} = \sum_{i \in N_{t-1}} w_{i,t-1} r_{i,t} - \sum_{i \in (N_{t-1} \cup N_{t-2})} |w_{i,t-1} - w_{i,t-2}| q_{i,t-1} \tag{12}$$

r_t^{AC} is the portfolio return after transaction costs on day t, N_{t-1} is the set of stocks in the portfolio over day t as determined at the end of day $t-1$, $w_{i,t-1}$ is the portfolio weight of stock i at time of portfolio formation at the end of day $t-1$, $r_{i,t}$ is the raw return from the end of day $t-1$ to the end of day t, and $q_{i,t-1}$ is the quoted half-spread at the end of the day $t-1$. Stocks that get delisted on day $t-1$ do not have a valid quoted-half spread on day t and we assume that these stocks earn the delisting return without further transaction costs. If a stock has a valid quoted half-spread on day $t-1$, but not on day t, we take the quoted half-spread from day $t-1$ to approximate the trading costs. This procedure is necessary if a stock no longer belongs to the tradeable universe according to our filters. We then close the position assuming that the spread of the previous day equals the spread today. The union in the second summand ensures that trading costs are also paid if a position is sold entirely.

To reduce trading costs, we apply a straight-forward trading cost mitigation strategy. On each day, we only consider stocks in the low-cost universe. Following Novy-Marx and

Velikov (2015), we restrict our strategy to those stocks that lie within the terctile with lowest bid-ask spreads at the time of portfolio formation.[13]

We calculate the portfolio value of the trading cost minimized long-only 4/5 strategy under the assumption that the strategy starts with an investment amount of $1m dollar on January 1, 2004. Between 2004 and in the financial crisis, the long 4/5 strategy performs approximately like the market. In this time period, excess returns are roughly eaten up by transaction costs implied by quoted bid-ask spreads. Spreads and returns before costs increase dramatically during the financial crisis, but the overall picture of excess returns being roughly equal to transaction costs remains unchanged.

However, the situation changes significantly after the financial crisis. Spreads come down to levels below the spreads we have seen in the years preceding the financial crisis, and continue to fall at low levels. At the end of the sample, the average of the value-weighted quoted half-spreads of portfolios 4 and 5 approach 1 basis point. At the same time, returns do not fall as much as spreads, making positive after-cost excess returns possible. After costs, the daily alpha of the strategy's excess return over the risk-free-rate with respect to the five Fama-French factors is 2.586 bps with a Newey and West (1987) adjusted t-statistic of 2.948 in the time period from January 1, 2004, until the end of the sample in December 2019. The strategy loads negatively on SMB (Small Minus Big), RMW (Robust Minus Weak), and CMA (Conservative Minus Aggressive).

The main take-away of the empirical exercise is that the dramatic increase in liquidity since the early 2000's opens the door for after-cost profitablity of the value-weighted portfolio strategy with daily rebalancing; in contrast to the situation in the twentieth century.[14] There

[13] Novy-Marx and Velikov (2015) perform a double sort on size and spreads "to avoid a large cap bias" (page 138). The goal of our exercise is to implement a strategy among the most liquid stocks, irrespective of size. We therefore only condition on liquidity each day, accepting that this will bias our universe of stocks towards large-caps.

[14] The main open question is the potential price impact of trades. There is reason to believe that streak strategies incur much smaller price impact costs than most other daily strategies. First, our main results are based on value-weighted portfolios, and most trading takes place in liquid stocks. Second, our strategy buys on days with falling prices and sells on days with rising prices, suggesting that a financial institution implementing the strategy is rather liquidity provider than demander. However, the reliable estimation of the strategy's capacity is not possible due to the lack of the necessary data.

are certainly further ways to increase profitability after transaction costs (Novy-Marx and Velikov, 2015). However, even if a short-term strategy is not profitable after trading costs, it is still possible that the reported return predictability provides useful information for optimizing already existing trading and/or transaction-costs-minimization strategies.[15]

5.4 Earnings Announcements

Recent evidence suggests that behavioral agents trade more aggressively during the days preceding an earnings announcement. Prior to an earnings announcement, retail demand for lottery-stocks (Liu, Wang, Yu, and Zhao, 2019) and for stocks with recent previous earnings surprises (Frieder, 2008, Shanthikumar, 2012, Ertan, Karolyi, Kelly, and Stoumbos, 2019) increases. The increased demand is associated with abnormal price increases before earnings announcements and reversals afterwards (Liu, Wang, Yu, and Zhao, 2019, Ertan, Karolyi, Kelly, and Stoumbos, 2019).[16] To the extent that these results carry over to return extrapolation, we would expect returns from streak strategies to be higher if the holding day coincides with an earnings announcement day. Initial evidence consistent with this prediction comes from So and Wang (2014), who document that an equally-weighted short-term reversal strategy based on cumulative three-day-returns is more profitable around earnings announcements.[17]

Table 10 reports the average returns on holding days with and without an earnings announcement. Averages are weighted by the stock's market capitalization on that day.

[15] Our results seem to be at odds with recent evidence reported by Chen and Velikov (2019). They show that many anomalies face high trading costs post-publication, that is, in recent years. The reason for this qualitative difference is that the long 4/5 strategy trades only the most liquid and largest stocks, while the anomalies analyzed by Chen and Velikov (2019) often form equally-weighted portfolios over the entire CRSP universe, leading to portfolios with much higher average spreads.

[16] Going beyond short-term horizons, several authors have investigated patterns in earning surprises (see, e.g., Barth, Elliott, and Finn, 1999), including streaks (Loh and Warachka, 2012), and their implications for longer term returns.

[17] So and Wang (2014) interpret their empirical result as an increased compensation for liquidity provision prior to uncertain information events. Our alternative, mutually non-exclusive interpretation is that extrapolators cause more mispricing prior to earnings announcements and that this mispricing gets eliminated afterwards.

Table 10: Streak portfolios' returns and earnings announcements

For both positive and negative streaks this table reports the average excess-market returns for different streak lengths on days with and without earnings announcements (EA-days and no-EA-days). Observations are weighted by the firm's market-capitalization relative to the total market capitalization of the whole market on that day. We report a weighted two-sample t-test to assess the difference in means.

	Negative streaks					Positive streaks				
	Length of streak (days)					Length of streak (days)				
	1	2	3	4	5	1	2	3	4	5
no-EA-days	0.006	0.036	0.056	0.089	0.105	-0.006	-0.030	-0.047	-0.057	-0.077
EA-days	0.140	0.219	0.229	0.385	0.410	-0.117	-0.191	-0.246	-0.208	-0.317
Difference	0.135	0.184	0.173	0.296	0.306	-0.112	-0.161	-0.198	-0.151	-0.240
t-stat	11.140	10.401	6.704	8.015	5.613	-9.412	-9.455	-8.016	-4.043	-4.191

Returns are substantially more extreme on earnings-announcement days and significantly different from streak returns on non-announcement days.

The results are broadly consistent with the idea that earnings announcements reduce mispricing caused by disagreement among market participants (Berkman, Dimitrov, Jain, Koch, and Tice, 2009, Engelberg, McLean, and Pontiff, 2018, Daniel, Klos, and Rottke, 2020a).[18]

[18] In Appendix D, we study the interaction between institutional ownership and streak returns. We find that the profitability of streak strategies is highest among stocks with low institutional ownership. This is broadly consistent with results reported by Da, Huang, and Jin (2019) and with previous studies on the relationship between mispricing and institutional ownership (see, for example Nagel, 2005). We further find that value-weighted portfolios returns from streak strategy are smaller for a portfolio composed of medium-IOR-stocks compared to a portfolio of high-IOR stocks. We do not discuss these results in the main text for two reasons. First, although there is direct evidence that some individual investors act as noise traders (Foucault, Sraer, and Thesmar, 2011), recent research has also shown that the group of retail investors is heterogeneous. Retail investors as a group appear to be sophisticated in the sense that their daily buy-sell-imbalance predicts returns on a monthly horizon (Kelley and Tetlock, 2013). Some retail investors even act like informed short sellers (Kelley and Tetlock, 2017). Second, recent results have raised doubts that the majority of institutions tries to arbitrage anomalies. Particularly, institutions as a group buy overvalued stocks as judged by commonly known anomalies and the change in institutional ownership before an earnings announcement is a negatively related to returns around earnings announcements (Edelen, Ince, and Kadlec, 2016). In sum, the precise relationship between the observed institutional ownership ratio and the unobserved amount of arbitrage capital that flows to a stock seems not to be fully understood.

6 Conclusion

A simplified version of the extrapolation model by Da, Huang, and Jin (2019) implies that streaks in returns have predictive power for future returns. Consistent with the theoretical predictions, we find that value-weighted long-short portfolios based on streaks in daily returns earn sizable returns in the U.S. and in international markets. Diversifying across different streak lengths and internationally yields annualized Sharpe ratios before costs of 3.25. Several additional tests in the U.S. show that the form of return predictability documented in this paper is economically different from previously reported forms of daily return predictability and cannot be explained by illiquidity.

References

Abdi, Farshid, and Angelo Ranaldo, 2017, A simple estimation of bid-ask spreads from daily close, high, and low prices, *Review of Financial Studies* 30, 4437–4480.

Amini, Shima, Bartosz Gebka, Robert Hudson, and Kevin Keasey, 2013, A review of the international literature on the short term predictability of stock prices conditional on large prior price changes: Microstructure, behavioral and risk related explanations, *International Review of Financial Analysis* 26, 1–17.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.

Asness, Clifford, Tobias Moskowitz, and Lasse Heje Pedersen, 2013, Value and momentum everywhere, *Journal of Finance* 68, 929–985.

Asparouhova, Elena, Hendrik Bessembinder, and Ivalina Kalcheva, 2013, Noisy prices and inference regarding returns, *Journal of Finance* 68, 665–714.

Atkins, Allen B., and Edward A. Dyl, 1990, Price reversals, bid-ask spreads, and market efficiency, *Journal of Financial and Quantitative Analysis* 25, 535–547.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427–446.

Barber, Brad M., and John D. Lyon, 1997, Detecting long-run abnormal stock returns: The empirical power and specification of test statistics, *Journal of Financial Economics* 43, 341–372.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2015, X-CAPM: An extrapolative capital asset pricing model, *Journal of Financial Economics* 115, 1–24.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2018, Extrapolation and bubbles, *Journal of Financial Economics* 129, 203–227.

Barclay, Michael J., William G. Christie, Jeffrey H. Harris, Eugene Kandel, and Paul H. Schultz, 1999, Effects of market reform on the trading costs and depths of NASDAQ stocks, *Journal of Finance* 54, 1–34.

Barth, Mary E. John A. Elliott, and Mark W. Finn, 1999, Market rewards associated with patterns of increasing earnings, *Journal of Accounting Research* 37, 387–413.

Berkman, Henk, Valentin Dimitrov, Prem C. Jain, Paul D. Koch, and Sheri Tice, 2009, Sell on the news: Differences of opinion, short-sales constraints, and returns around earnings announcements, *Journal of Financial Economics* 92, 376–399.

Bessembinder, Hendrik, 2003, Trade execution costs and market quality after decimalization, *Journal of Financial and Quantitative Analysis* 38, 747–777.

Bessembinder, Hendrik, and Kumar Venkataraman, 2010, Bid–ask spreads, in Rama Cont, ed., *Encyclopedia of Quantitative Finance*, 184–189 (John Wiley & Sons).

Blake, Andrew, and Andrew Zisserman, 1987, *Visual Reconstruction* (MIT Press, Cambridge, MA, USA).

Blume, Marshall E., and Robert F. Stambaugh, 1983, Biases in computed returns: An application to the size effect, *Journal of Financial Economics* 12, 387–404.

Boudoukh, Jacob, Matthew P. Richardson, and Robert F. Whitelaw, 1994, A tale of three schools: Insights on autocorrelations of short-horizon stock returns, *Review of Financial Studies* 7, 539–573.

Bremer, Marc, and Richard J. Sweeney, 1991, The reversal of large stock-price decreases, *Journal of Finance* 46, 747–754.

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Chen, Andrew Y., and Mihail Velikov, 2019, Accounting for the anomaly zoo: A trading cost perspective, Working Paper.

Chordia, Tarun, Avanidhar Subrahmanyam, and Qing Tong, 2014, Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?, *Journal of Accounting and Economics* 58, 41–58.

Chung, Kee H., and Hao Zhang, 2014, A simple approximation of intraday spreads using daily data, *Journal of Financial Markets* 17, 94–120.

Coibion, Olivier, and Yuriy Gorodnichenko, 2015, Information rigidity and the expectations formation process: A simple framework and new facts, *American Economic Review* 105, 2644–2678.

Comerton-Forde, Carole, and Tālis J. Putniņš, 2011, Measuring closing price manipulation, *Journal of Financial Intermediation* 20, 135–158.

Cox, Don R., and David R. Peterson, 1994, Stock returns following large one-day declines: Evidence on short-term reversals and longer-term performance, *Journal of Finance* 49, 255–267.

Da, Zhi, Xing Huang, and Lawrence Jin, 2019, Extrapolative beliefs in the cross-section: What can we learn from the crowds?, Working Paper.

Daniel, Kent, Alexander Klos, and Simon Rottke, 2020a, Overconfidence, information diffusion, and mispricing persistence, Working Paper.

Daniel, Kent, Lira Mota, Simon Rottke, and Tano Santos, 2020b, The cross-section of risk and return, *Review of Financial Studies* 33, 1927–1979.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

De Groot, Wilma, Joop Huij, and Weili Zhou, 2012, Another look at trading costs and short-term reversal profits, *Journal of Banking and Finance* 36, 371–382.

Duffie, Darrell, 2010, Asset price dynamics with slow-moving capital, *Journal of Finance* 65, 1237–1267.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.

Edelen, Roger M., Ozgur S. Ince, and Gregory B. Kadlec, 2016, Institutional investors and stock return anomalies, *Journal of Financial Economics* 119, 472–488.

Engelberg, Joseph, R. David McLean, and Jeffrey Pontiff, 2018, Anomalies and news, *Journal of Finance* 73, 1971–2001.

Engle, Robert, Robert Ferstenberg, and Jeffrey Russell, 2012, Measuring and modeling execution cost and risk, *Journal of Portfolio Management* 38, 14–28.

Ertan, Aytekin, Stephen Karolyi, Peter Kelly, and Robert Stoumbos, 2019, Earnings announcement return extrapolation, Working Paper.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2012, Size, value, and momentum in international stock returns, *Journal of Financial Economics* 105, 457–472.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1 – 22.

Fama, Eugene F. and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Foucault, Thierry, David Alexandre Sraer, and David J. Thesmar, 2011, Individual investors and volatility, *Journal of Finance* 66, 1369–1406.

Frazzini, Andrea, Ronen Israel, and Tobias Moskowitz, 2012, Trading costs of asset pricing anomalies, Working Paper.

Frazzini, Andrea, Ronen Israel, and Tobias Moskowitz, 2018, Trading costs, Working Paper.

Frieder, Laura, 2008, Investor and price response to patterns in earnings surprises, *Journal of Financial Markets* 11, 259–283.

Green, Jeremiah, John R. M. Hand, and X. Frank Zhang, 2017, The characteristics that provide independent information about average U.S. monthly stock returns, *Review of Financial Studies* 30, 4389–4436.

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of returns and expected returns, *Review of Financial Studies* 27, 714–746.

Griffin, John M., Patrick J. Kelly, and Federico Nardari, 2010, Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets, *Review of Financial Studies* 23, 3225–3277.

Gutierrez, Roberto C., and Eric K. Kelley, 2008, The long-lasting momentum in weekly returns, *Journal of Finance* 63, 415–447.

Hameed, Allaudeen, Wenjin Kang, and S. Viswanathan, 2010, Stock market declines and liquidity, *Journal of Finance* 25, 747–777.

Hendershott, Terrence, Charles M. Jones, and Albert Menkveld, 2011, Does algorithmic trading improve liquidity?, *Journal of Finance* 66, 1–33.

Hou, Kewei, Chen Xue, and Lu Zhang, 2019, Replicating anomalies, *Review of Financial Studies* forthcoming.

Ince, Ozgur S., and R. Burt Porter, 2006, Individual equity return data from Thomson Datastream: Handle with care!, *Journal of Financial Research* 29, 463–479.

Jacobs, Heiko, 2016, Market maturity and mispricing, *Journal of Financial Economics* 122, 270–287.

Jacobs, Heiko, and Sebastian Müller, 2020, Anomalies across the globe: Once public, no longer existent?, *Journal of Financial Economics* 135, 213–230.

Jiang, George J., and Kevin X. Zhu, 2017, Information shocks and short-term market underreaction, *Journal of Financial Economics* 124, 43–64.

Keim, Donald B., and Ananth Madhavan, 1997, Transactions costs and investment style: An inter-exchange analysis of institutional equity trades, *Journal of Financial Economics* 46, 265 – 292.

Kelley, Eric K., and Paul C. Tetlock, 2013, How wise are crowds? insights from retail orders and stock returns, *Journal of Finance* 68, 1229–1265.

Kelley, Eric K., and Paul C. Tetlock, 2017, Retail short selling and stock prices, *Review of Financial Studies* 30, 801–834.

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs?, *Journal of Finance* 59, 1039–1082.

Kumar, Alok, Stefan Ruenzi, and Michael Ungeheuer, 2019, Daily winners and losers, Working Paper.

Lesmond, David A., Michael J. Schill, and Chunsheng Zhou, 2004, The illusory nature of momentum profits, *Journal of Financial Economics* 71, 349–380.

Liu, Bibo, Huijun Wang, Jianfeng Yu, and Shen Zhao, 2019, Time-varying demand for lottery: Speculation ahead of earnings announcements, Working Paper.

Loh, Roger K. and Mitch Warachka, 2012, Streaks in earnings surprises and the cross-section of stock returns, *Management Science* 58, 1305–1321.

MacKinlay, A Craig, 1997, Event studies in economics and finance, *Journal of Economic Literature* 35, 13–39.

Mankiw, N. Gregory, and Ricardo Reis, 2002, Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve, *Quarterly Journal of Economics* 117, 1295–1328.

Merton, Robert C., 1976, Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics* 3, 125–144.

Nagel, Stefan, 2005, Short sales, institutional investors and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309.

Nagel, Stefan, 2012, Evaporating liquidity, *Review of Financial Studies* 25, 2005–2039.

Newey, Whitney, and Kenneth West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–08.

Novy-Marx, Robert, and Mihail Velikov, 2015, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

Savor, Pavel G., 2012, Stock returns after major price shocks: The impact of information, *Journal of Financial Economics* 106, 635–659.

Schmidt, Peter, Urs Von Arx, Andreas Schrimpf, Alexander Wagner, and Andreas Ziegler, 2017, On the construction of common size, value and momentum factors in international stock markets: A guide with applications, Working Paper.

Shanthikumar, Devin M., 2012, Consecutive earnings surprises: Small and large trader reactions, *The Accounting Review* 87, 1709–1736.

So, Eric C. and Sean Wang, 2014, News-driven return reversals: Liquidity provision ahead of earnings announcements, *Journal of Financial Economics* 114, 20–35.

Tetlock, Paul C., 2011, All the news that's fit to reprint: Do investors react to stale information?, *Review of Financial Studies* 24, 1481–1512.

Appendix

A Proof

Proof of Proposition 1. To derive the corresponding formula, it is useful to take a closer look at the early price changes in an economy that sees it first dividend innovation ϵ_1 in period 1. Da, Huang, and Jin (2019) call $\frac{1}{1-(\mu^E/\mu^F)\lambda_1(1-\lambda_2)}$ the amplification factor. For the sake of brevity, we set $A = (\mu^E/\mu^F)\lambda_1(1-\lambda_2)$ and therefore $\frac{1}{1-A} = \frac{1}{1-(\mu^E/\mu^F)\lambda_1(1-\lambda_2)}$. For the first price change, we have:

$$P_1 - P_0 = \frac{1}{1-A}\left[\epsilon_1 + \gamma\sigma^2 Q\right] - \frac{A}{(1-A)(1-\lambda_2)}(1-\lambda_2)S_0 \tag{A.1}$$

(We deliberately do not cancel the $(1-\lambda_2)$ in the second term.) The second price change can be written as

$$\begin{aligned}
P_2 - P_1 &= \frac{1}{1-A}\left[\epsilon_2 + \gamma\sigma^2 Q\right] - \frac{A}{(1-A)(1-\lambda_2)}(\lambda_2 - \lambda_2^2)S_0 \\
&\quad - \frac{A}{(1-A)}\left[(1-\lambda_2)\left[\frac{1}{1-A}\left[\epsilon_1 + \gamma\sigma^2 Q\right] - \frac{A}{(1-A)(1-\lambda_2)}(1-\lambda_2)S_0\right]\right] \\
&= \frac{1}{1-A}\left[\epsilon_2 + \gamma\sigma^2 Q\right] - \frac{A}{(1-A)^2}(1-\lambda_2)\left[\epsilon_1 + \gamma\sigma^2 Q\right] \\
&\quad - \frac{A}{(1-A)(1-\lambda_2)}(\lambda_2 - \lambda_2^2)S_0 + \frac{A^2}{(1-A)^2(1-\lambda_2)}(1-\lambda_2)^2 S_0
\end{aligned} \tag{A.2}$$

Now consider the coefficients of the summands who contain the dividend innovation ϵ_1. In $P_2 - P_1$, ϵ_1 appears in just one summand with coefficient $-\frac{A}{(1-A)^2}(1-\lambda_2)$. ϵ_1 appears one-time in $P_2 - P_1$ because $P_1 - P_0$ enters $P_2 - P_1$ one-time with coefficient $-\frac{A}{(1-A)}(1-\lambda_2)$ and ϵ_1 enters $P_1 - P_0$ one-time with coefficient $\frac{1}{1-A}$.

The third price change can be written as

$$\begin{aligned}
P_3 - P_2 &= \frac{1}{1-A}\left[\epsilon_3 + \gamma\sigma^2 Q\right] - \frac{A}{(1-A)(1-\lambda_2)}(\lambda_2^2 - \lambda_2^3)S_0 \\
&\quad - \frac{A}{(1-A)}\left[(1-\lambda_2)(P_2 - P_1) + (\lambda_2 - \lambda_2^2)(P_1 - P_0)\right] \\
&= \frac{1}{1-A}\left[\epsilon_3 + \gamma\sigma^2 Q\right] - \frac{A}{(1-A)^2}(1-\lambda_2)\left[\epsilon_2 + \gamma\sigma^2 Q\right] \\
&\quad + \frac{A^2}{(1-A)^3}(1-\lambda_2)^2\left[\epsilon_1 + \gamma\sigma^2 Q\right] - \frac{A}{(1-A)^2}(\lambda_2 - \lambda_2^2)\left[\epsilon_1 + \gamma\sigma^2 Q\right] \\
&\quad - \frac{A}{(1-A)(1-\lambda_2)}(\lambda_2^2 - \lambda_2^3)S_0 \\
&\quad + \frac{A^2}{(1-A)^2(1-\lambda_2)}(1-\lambda_2)(\lambda_2 - \lambda_2^2)S_0 - \frac{A^3}{(1-A)^3(1-\lambda_2)}(1-\lambda_2)^3 S_0
\end{aligned}$$

$$+\frac{A^2}{(1-A^2)(1-\lambda_2)}(1-\lambda_2)(\lambda_2-\lambda_2^2)S_0 \tag{A.3}$$

In $P_3 - P_2$, ϵ_1 now appears in two summands with coefficients $\frac{A^2}{(1-A)^3}(1-\lambda_2)^2$ and $-\frac{A}{(1-A)^2}(\lambda_2-\lambda_2^2)$.

1. The coefficient of the first summand with factor $(\epsilon_1+\gamma\sigma^2 Q)$ comes from the fact that ϵ_1 enters $P_2 - P_1$ one time with coefficient $-\frac{A}{(1-A)^2}(1-\lambda_2)$. $P_2 - P_1$ has now coefficient $-\frac{A}{(1-A)}(1-\lambda_2)$ in $P_3 - P_2$. The overall coefficient of the first appearance of ϵ_1 must therefore be $\frac{A^2}{(1-A)^3}(1-\lambda_2)^2$.

2. The coefficient of the second summand with factor $(\epsilon_1+\gamma\sigma^2 Q)$ comes from the fact that ϵ_1 enters $P_1 - P_0$ one time with coefficient $-\frac{1}{(1-A)}$. $P_1 - P_0$ has coefficient $-\frac{A}{(1-A)}(\lambda_2-\lambda_2^2)$ in $P_3 - P_2$. The overall coefficient of the second appearance of ϵ_1 must therefore be $\frac{A}{(1-A)^2}(\lambda_2-\lambda_2^2)$.

There is a helpful way to think about these two changes that have happened to the original coefficient of ϵ_1 in $P_2 - P_1$ as we build $P_3 - P_2$.

1. Reproduction: The ϵ_1-term with the old coefficient $-\frac{A}{(1-A)^2}(1-\lambda_2)$ already incorporated in $P_2 - P_1$ reproduces itself as this coefficient is also part of $P_3 - P_2$. The new coefficient is the old coefficient $-\frac{A}{(1-A)^2}(1-\lambda_2)$ times $-\frac{A}{(1-A)}(1-\lambda_2)$.

2. Aging: The ϵ_1-term with the old coefficient $-\frac{A}{(1-A)^2}(1-\lambda_2)$ already incorporated in $P_2 - P_1$ is also still part of $P_3 - P_2$ as $P_3 - P_2$ is a function of $P_2 - P_1$. However, the coefficient becomes "older" in the sense, that he receives weight $(\lambda_2 - \lambda_2^2)$ instead of $(1-\lambda_2)$, thereby "transforming" $-\frac{A}{(1-A)^2}(1-\lambda_2)$ into $-\frac{A}{(1-A)^2}(\lambda_2-\lambda_2^2)$.

"Reproduction" and "Aging" happens now to all coefficients when we go one price change in the future, i.e. from $P_t - P_{t-1}$ to $P_{t+1} - P_t$. Note that any "Reproduction" comes along with a sign flip of the coefficient. Figure A.1 illustrates this process for summands with the term $(\epsilon_1+\gamma\sigma^2 Q)$ for the first four price changes.

Generalizing this reasoning, we can write all summands that include the factor $(\epsilon_1+\gamma\sigma^2 Q)$ in price difference $P_t - P_{t-1}$ as

$$\frac{A}{(1-A)^2}\left[\sum_{i=0}^{t-2}\binom{t-2}{i}\left(\frac{A}{1-A}\right)^i(1-\lambda_2)^i(-1)^{i+1}\left(\lambda_2^{t-2-i}-\lambda_2^{t-1-i}\right)\right](\epsilon_1+\gamma\sigma^2 Q) \tag{A.4}$$

The summands that include the factor $(\epsilon_2+\gamma\sigma^2 Q)$ in price difference $P_t - P_{t-1}$ evolve exactly in the same way, except that the dividend innovation ϵ_2 starts one period later than ϵ_1. We can write all summands as

$$\frac{A}{(1-A)^2}\left[\sum_{i=0}^{t-3}\binom{t-3}{i}\left(\frac{A}{1-A}\right)^i(1-\lambda_2)^i(-1)^{i+1}\left(\lambda_2^{t-3-i}-\lambda_2^{t-2-i}\right)\right](\epsilon_2+\gamma\sigma^2 Q) \tag{A.5}$$

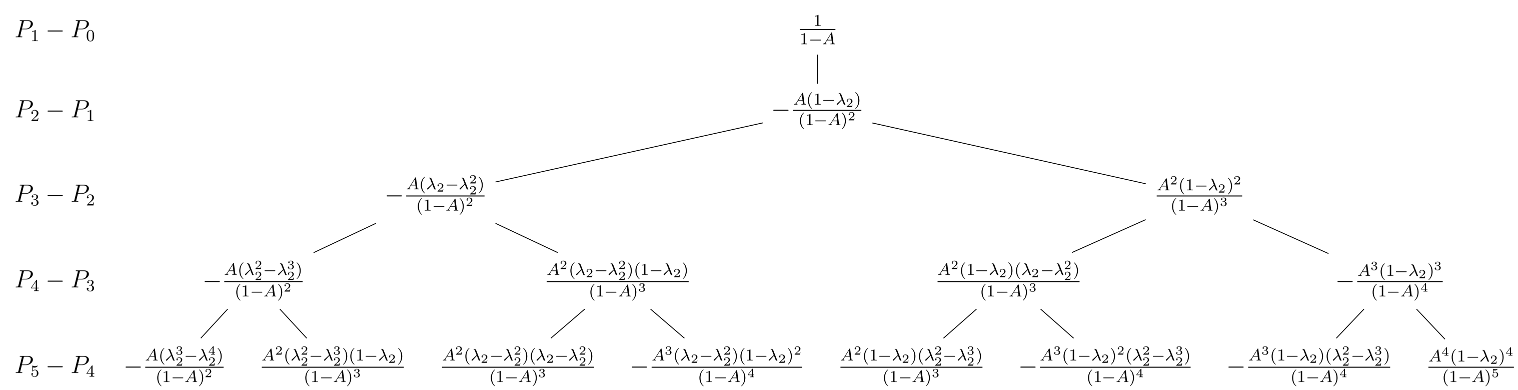

Figure A.1: Coefficients of summands including the factor $(\epsilon_1 + \gamma\sigma^2 Q)$ in the first four price differences

Applying this logic to all t dividend innovations that determine $P_t - P_{t-1}$ yields a formula for all summands in price difference $P_t - P_{t-1}$ that include a dividend innovation:

$$\left(\frac{1}{1-A}\right)(\epsilon_t + \gamma\sigma^2 Q) + \\ \sum_{j=1}^{t-1}(\epsilon_j + \gamma\sigma^2 Q)\left[\frac{A}{(1-A)^2}\sum_{i=0}^{t-1-j}\binom{t-1-j}{i}\left(\frac{A}{1-A}\right)^i(1-\lambda_2)^i(-1)^{i+1}\left(\lambda_2^{t-1-j-i} - \lambda_2^{t-j-i}\right)\right] \tag{A.6}$$

Note that $(-1)^x = (-1)^{-x}\ \forall x \in \mathbb{Z}$. We can write equation (A.6) as

$$\left(\frac{1}{1-A}\right)(\epsilon_t + \gamma\sigma^2 Q) + \\ \sum_{j=1}^{t-1}(\epsilon_j + \gamma\sigma^2 Q)\Bigg[\frac{A}{(1-A)^2}\sum_{i=0}^{t-1-j}\binom{t-1-j}{i}\left(\frac{A}{1-A}(1-\lambda_2)\right)^i(-1)^{t-j}(-\lambda_2)^{t-1-j-i} \\ - \frac{A\lambda_2}{(1-A)^2}\sum_{i=0}^{t-1-j}\binom{t-1-j}{i}\left(\frac{A}{1-A}(1-\lambda_2)\right)^i(-1)^{t-j}(-\lambda_2)^{t-1-j-i}\Bigg] \tag{A.7}$$

The general binomial theorem allows us to simplify to

$$\left(\frac{1}{1-A}\right)(\epsilon_t + \gamma\sigma^2 Q) + \\ \sum_{j=1}^{t-1}(\epsilon_j + \gamma\sigma^2 Q)\left[\left(\frac{A}{(1-A)^2} - \frac{A\lambda_2}{(1-A)^2}\right)\left(\frac{A}{1-A}(1-\lambda_2) - \lambda_2\right)^{t-1-j}(-1)^{t-j}\right] \tag{A.8}$$

and finally

$$\left(\frac{1}{1-A}\right)(\epsilon_t + \gamma\sigma^2 Q) + \frac{A}{(1-A)^2}(1-\lambda_2)\sum_{j=1}^{t-1}(\epsilon_j + \gamma\sigma^2 Q)\left(\frac{A}{1-A}(1-\lambda_2) - \lambda_2\right)^{t-1-j}(-1)^{t-j} \tag{A.9}$$

We can make a similar argument with regard to the coefficients of all terms that share S_0 as a factor. All terms including S_0 as a factor can be written as

$$-\frac{A}{(1-A)(1-\lambda_2)}(\lambda_2^{t-1} - \lambda_2^t)S_0 - \left(\frac{A}{1-A}\right)^2 S_0 \sum_{j=1}^{t-1}(\lambda_2^{j-1} - \lambda_2^j)\left(\frac{A}{1-A}(1-\lambda_2) - \lambda_2\right)^{t-1-j}(-1)^{t-j} \tag{A.10}$$

Combining the terms from equations (A.9) and (A.10), we get

$$P_t - P_{t-1} =$$
$$\left(\frac{1}{1-A}\right)(\epsilon_t + \gamma\sigma^2 Q) + \frac{A}{(1-A)^2}(1-\lambda_2)\sum_{j=1}^{t-1}(\epsilon_j + \gamma\sigma^2 Q)(-1)^{t-j}\left(\frac{A}{1-A}(1-\lambda_2) - \lambda_2\right)^{t-1-j}$$
$$- \frac{A}{1-A}\lambda_2^{t-1}S_0 - \left(\frac{A}{1-A}\right)^2 S_0 \sum_{j=1}^{t-1}\lambda_2^{j-1}(1-\lambda_2)(-1)^{t-j}\left(\frac{A}{1-A}(1-\lambda_2) - \lambda_2\right)^{t-1-j} \tag{A.11}$$

We now assume that extrapolators are unbiased at $t = 0$ in the sense that their expectation of the first price change is equal to the market risk premium if sentiment directly determines demand ($\lambda_0 = 0$ and $\lambda_1 = 1$). This assumption implies $S_0 = \gamma\sigma^2 Q$.[19] All terms with the market risk premium $\gamma\sigma^2 Q$ are given by

$$\begin{aligned} &\left(\frac{1}{1-A}\right)\gamma\sigma^2 Q - \frac{A}{(1-A)^2}(1-\lambda_2)\gamma\sigma^2 Q\sum_{j=1}^{t-1}\left(\lambda_2 - \frac{A}{1-A}(1-\lambda_2)\right)^{t-1-j} \\ &- \frac{A}{1-A}\lambda_2^{t-1}\gamma\sigma^2 Q + \left(\frac{A}{1-A}\right)^2(1-\lambda_2)\gamma\sigma^2 Q\sum_{j=1}^{t-1}\lambda_2^{j-1}\left(\lambda_2 - \frac{A}{1-A}(1-\lambda_2)\right)^{t-1-j} \end{aligned} \tag{A.12}$$

Equation (A.12) contains the sums of two finite geometric series.
Note that $\sum_{j=1}^{t-1} x^{t-1-j} = x^{t-1}\left[\sum_{j=0}^{t-1} x^{-j} - 1\right] = \frac{x^{t-1}-1}{x-1}$ and
$\sum_{j=1}^{t-1}\lambda_2^{j-1}x^{t-1-j} = \frac{x^{t-1}}{\lambda_2}\left[\sum_{j=0}^{t-1}\left(\frac{\lambda_2}{x}\right)^j - 1\right] = \frac{x^{t-1}\lambda_2 - \lambda_2^t}{\lambda_2(x-\lambda_2)}$ with $x = \lambda_2 - \frac{A}{1-A}(1-\lambda_2) = \frac{\lambda_2 - A}{1-A}$.
Using these observations to rewrite equation (A.12) yields

$$\begin{aligned} &\left(\frac{1}{1-A}\right)\gamma\sigma^2 Q - \frac{A}{(1-A)^2}(1-\lambda_2)\gamma\sigma^2 Q\frac{\left(\frac{\lambda_2-A}{1-A}\right)^{t-1} - 1}{\frac{\lambda_2-A}{1-A} - 1} \\ &- \frac{A}{1-A}\lambda_2^{t-1}\gamma\sigma^2 Q + \left(\frac{A}{1-A}\right)^2(1-\lambda_2)\gamma\sigma^2 Q\frac{\left(\frac{\lambda_2-A}{1-A}\right)^{t-1}\lambda_2 - \lambda_2^t}{\lambda_2\left(\frac{\lambda_2-A}{1-A} - \lambda_2\right)} \end{aligned} \tag{A.13}$$

After some tedious rearrangements, equation (A.13) evaluates to just $\gamma\sigma^2 Q$. Using this insight together with equation (A.11) gives the formula in the proposition.

□

[19] Barberis, Greenwood, Jin, and Shleifer (2018) use the same assumption for initial sentiment.

B Simulation of Cross-Sectional Extrapolation Model

We simulate asset prices for a market with 1,000 assets and 100 time periods in order to illustrate the theoretical results and economic intuitions of the model. We allow each firm i to have a different pair of extrapolation parameters $\lambda_{i,1}$ and $\lambda_{i,2}$. To emphasize this fact, we will keep the index i, in contrast to large parts of the main text.

We set $\gamma = 0.01$ and $\mu^E = 0.5$. $\lambda_{i,1}$ is sampled from a normal distribution with mean 1 and a standard deviation of 1. We truncate the realizations of $\lambda_{i,1}$ in the sense that the realization cannot be smaller than 0.5 or larger than 5. Furthermore, $\lambda_{i,2}$ is sampled from a normal distribution with mean 0.7 and standard deviation 0.2 and required to be larger than 0 and smaller than 1. If $(\mu^E/\mu^F)\lambda_{i,1}(1 - \lambda_{i,2}) > 1$, we throw away the observation and draw another $\lambda_{i,1}$ and $\lambda_{i,2}$ observation.

In Table B.1, Fama and MacBeth (1973) regressions show the relationship between extrapolators' current sentiment and next period price changes. Consistent with the theoretical predictions of the model, the coefficient of this period sentiment is negative, showing that the next period price changes will be lower if the sentiment today is more positive and vice versa.

Table B.1: Relationship between sentiment and next day price change

Fama-Macbeth regression reporting the relationship between price change in t and sentiment in t-1. Price changes and sentiment are calculated based on the Da, Huang, and Jin (2019).

	$Price\ Change_t$
$Sentiment_{t-1}$	-0.242^{***}
	(-34.596)
Constant	0.398^{***}
	(32.790)

Our goal is to apply this insight to the entire cross-section of equity returns. The fundamental problem is that sentiment S_{it} is not directly observable and, as shown by equation (2), depends not only on past price changes, but also on the stock-specific, unobservable, and potentially time-varying parameter $\lambda_{i,2}$, not to mention the fact that further determinants of extrapolators' demands, $\lambda_{i,0}$ and $\lambda_{i,1}$, are not directly observable either and their time-series properties are unknown.

To circumvent these problems, we look for empirical proxies that are easy to use and correlate with sentiment $S_{i,t}$ in a reliable way. We start with equation (2), which defines $S_{i,t}$. Sentiment for stock i in period t is a non-linear function of recent price changes. If past price changes have opposite signs and $\lambda_{i,2}$ is not known, it is almost always hard to determine for the econometrician if sentiment on any given day is positive or negative. However, if all the most recent price changes have the same sign, sentiment today has this sign too, as long as

$\lambda_{i,2}$ and the absolute values of past price changes before the streak are not implausibly high in absolute terms. We therefore hypothesize that past price change streaks, i.e., several days of price changes of the same sign, are a simple and powerful predictor of the next day's price change. Section 2 in the main text develops this intuition in a rigorous way.

Table B.2 reports the results of the Fama-MacBeth regression with the simulated data. We observe that an increase in streak length of the most recent price changes increases the magnitude of the sentiment prevailing in the current period. This confirms that streak length is a good proxy for the extrapolators' sentiment on a given day. The relationship is monotonic, consistent with the intuition that sentiment tends to be higher the longer a price change streak lasted.

The model of Da, Huang, and Jin (2019) predicts further that extreme sentiment tends to be followed by price changes of the opposite sign, as sentiment comes back to non-extreme values in the absence of further extreme fundamental shocks. In Table B.3, we test whether price change streaks can be used as a proxy for sentiment to predict future price changes. Fama-MacBeth regressions show exactly this effect, since longer price change streaks increase the negative predictability of next period's price changes steadily, for streaks with both negative and positive price changes. Further, we see that the streak dummies lose significance when we include the current sentiment $S_{i,t}$, confirming that price change streaks per se do not convey relevant information. Rather, streaks serve as easy observable proxies for the current sentiment of extrapolators in our simulated data.

Table B.2: Sentiment and streak length

Fama-MacBeth regression testing the relationship between sentiment on the last period of a streak and the streak length are shown. The variable $Streak_5^+$ takes on the value 1 if the price changes of a given asset in the past 5 periods have been positive. The index $t-1$ highlights that the streaks are based on past price changes. The streak variables are constructed in such a manner that, when a stock has had 5 periods of positive price changes in the past, $Streak_5^+$ is one, while $Streak_4^+$, $Streak_3^+$, and $Streak_2^+$ are zero. The variable $Streak_5^-$ takes on the value 1 if the price changes of a given asset in the past 5 periods have been positive. To avoid the dummy trap caused by perfect multicollinearity in the streak dummy variables, the $Streak_{1,t-1}^+$ variable is removed from the specification in the third column.

		$Sentiment_t$	
$Streak_{1,t-1}^+$	0.195***		
	48.263		
$Streak_{1,t-2}^+$	0.301***		0.105***
	(50.904)		(16.661)
$Streak_{1,t-3}^+$	0.386***		0.190***
	(47.956)		(22.932)
$Streak_{1,t-4}^+$	0.427***		0.232***
	(40.211)		(20.954)
$Streak_{1,t-5}^+$	0.491***		0.296***
	(42.895)		(25.237)
$Streak_{1,t-1}^-$		−0.191***	−0.115***
		(−44.982)	(−24.462)
$Streak_{1,t-2}^-$		−0.309***	−0.233***
		(−60.494)	(−41.485)
$Streak_{1,t-3}^-$		−0.387***	−0.311***
		(−45.053)	(−35.856)
$Streak_{1,t-4}^-$		−0.452***	−0.376***
		(−36.147)	(−30.078)
$Streak_{1,t-5}^-$		−0.509***	−0.433***
		(−44.653)	(−35.999)
Constant	1.524***	1.795***	1.719***
	(157.939)	(181.097)	(177.780)

Table B.3: Relationship between price change, sentiment, and price change streak

Fama-MacBeth regressions with simulated price changes, $Price\ change_t$, as dependent variable are shown. The variable $Streak^+_{5,t-1}$ takes on the value 1 if the price changes of a given asset in the past 5 days have been positive. The index $t-1$ highlights that the streaks are based on past price changes. The streak variables are constructed in such a manner that, when a stock has had 5 days of positive price changes in the past, $Streak^+_{5,t-1}$ is one, while $Streak^+_{4,t-1}$, $Streak^+_{3,t-1}$, and $Streak^+_{2,t-1}$ are zero. The variable $Streak^-_5$ takes on the value 1 if the price changes of a given asset in the past 5 days have been positive. As in Da, Huang, and Jin (2019), $Sentiment_{t-1}$ is the sentiment on the last day of the price change streak. To avoid the dummy trap caused by perfect multicollinearity in the streak dummy variables, the $Streak^+_{1,t-1}$ variable is removed from the two specifications which includes the positive and negative *Streak* dummy variables.

			$Price\ change_t$			
$Streak^+_{1,t-1}$	-0.142^{***}		-0.054	-0.037^{***}		-0.054^{*}
	(-14.598)		(-1.626)	(-3.561)		(-1.897)
$Streak^+_{1,t-2}$	-0.214^{***}		-0.127^{***}	-0.053^{***}		-0.071^{**}
	(-17.047)		(-3.663)	(-3.856)		(-2.346)
$Streak^+_{1,t-3}$	-0.206^{***}		-0.118^{***}	-0.019		-0.038
	(-11.926)		(-3.243)	(-1.037)		(-1.146)
$Streak^+_{1,t-4}$	-0.243^{***}		-0.155^{***}	-0.039		-0.057
	(-9.756)		(-3.802)	(-1.495)		(-1.529)
$Streak^+_{1,t-5}$	-0.216^{***}		-0.129^{***}	-0.001		-0.020
	(-8.377)		(-3.164)	(-0.031)		$(-0.513\ )$
$Streak^-_{1,t-1}$		0.138^{***}	0.050		0.031^{***}	-0.025
		(14.163)	(1.507)		(2.997)	(-0.887)
$Streak^-_{1,t-2}$		0.210^{***}	0.122^{***}		0.054^{***}	-0.002
		(16.527)	(3.522)		(3.969)	(-0.060)
$Streak^-_{1,t-3}$		0.223^{***}	0.135^{***}		0.040^{**}	-0.016
		(12.422)	(3.679)		(2.142)	(-0.506)
$Streak^-_{1,t-4}$		0.238^{***}	0.151^{***}		0.038	-0.018
		(9.289)	(3.659)		(1.437)	(-0.486)
$Streak^-_{1,t-5}$		0.273^{***}	0.186^{***}		0.056^{**}	
		(10.164)	(4.419)		(1.960)	
$Sentiment_{t-1}$				-0.227^{***}	-0.223^{***}	-0.224^{***}
				(-27.599)	(-27.394)	(-26.466)
Constant	0.088^{***}	-0.087^{***}	0.000	0.391^{***}	0.347^{***}	0.404^{***}
	16.641	-16.431	0.000	31.612	20.727	14.492

Table B.4: Price change after positive and negative streaks in price changes

For the streak portfolio with different streak lengths of positive and negative price changes, this table reports portfolio price changes relative to the average market price change, standard deviations, t-stats, and average number of assets in the portfolio. The length of a streak is measured in number of periods ranging from 1 to 5. The values reported are those recorded in the period following the streak. The portfolios are formed using a simulated price process from the model by Da, Huang, and Jin (2019). The parameters used are the following $\gamma = 0.01$ and $\mu^E = 0.5$. $\lambda_{i,1}$ is sampled from a normal distribution with mean 1, a standard deviation of 1, and the condition that $\lambda_{i,1}$ cannot be smaller than 0.5 or larger than 5. Furthermore, $\lambda_{i,2}$ is sampled from a normal distribution, with mean 0.7, standard deviation 0.2, and conditioned to be larger than 0 and smaller than 1. The values of $\lambda_{i,1}$ and $\lambda_{i,2}$ need to fulfil the condition that $(\mu^E/\mu^F)\lambda_{i,1}(1 - \lambda_{i,2})$ is strictly smaller 1. The t-statistics are Newey-West t-statistics corrected for serial correlation and heteroskedasticity in the error term.

	Length of streak				
	1	2	3	4	5
Price changes after positive streaks	-0.091	-0.129	-0.132	-0.143	-0.149
Std. dev.	0.003	0.006	0.008	0.015	0.023
t-stat	-30.816	-21.526	-16.623	-9.623	-5.295
No. of risky assets	502	237.700	110.500	51.400	23.700
Price changes after negative streaks	0.092	0.131	0.147	0.166	0.177
Std. dev.	0.003	0.007	0.012	0.020	0.026
t-stat	29.416	19.041	12.476	8.304	6.920
No. of risky assets	498	233.800	106.200	47.800	21.500

C FF3-Alphas

Table C.1: Times series FF3-Alphas of streak portfolios

For the streak portfolios with different streak lengths, this table reports the alphas for the CAPM, the Fama-French three-factor model, the Fama-French five-factor model, and the associated Newey-West t-statistic. The length of a streak is measured in number of days ranging from 1 to 5.

		After negative streaks					After positive streaks				
Model	Stat	1	2	3	4	5	1	2	3	4	5
CAPM	α	0.007	0.038	0.059	0.094	0.110	-0.007	-0.033	-0.048	-0.056	-0.077
	t-stat	1.636	5.841	6.388	6.779	6.689	-1.757	-5.452	-5.776	-5.294	-5.349
FF3	α	0.007	0.039	0.059	0.094	0.110	-0.007	-0.033	-0.049	-0.057	-0.078
	t-stat	1.710	5.557	5.942	7.834	7.141	-1.839	-4.306	-4.877	-5.150	-5.328
FF5	α	0.009	0.043	0.066	0.103	0.122	-0.008	-0.035	-0.050	-0.059	-0.080
	t-stat	2.206	6.613	6.858	8.581	6.713	-1.996	-4.596	-5.190	-5.273	-5.375

D Institutional Ownership in Streak Stocks

Da, Huang, and Jin (2019) use IOR as a proxy to determine the share of extrapolators involved in a stock. More precisely, low IOR would indicate a higher share of extrapolators and vice versa. Based on the literature, we expect the holding day returns after streaks to be extremer when the ratio of stocks held by institutions is lower, because the mispricing is assumed to be higher with a lower level of non-institutional owners (Nagel, 2005).

We compute institutional ownership based on 13-F filings provided by Thomson-Reuters and WRDS SEC data.[20] The institutional ownership ratio (IOR) is the number of shares held by institutional owners divided by the number of shares outstanding reported in CRSP. As in Nagel (2005), the IOR of stocks without reported institutional holdings is set to zero. Institutional ownership is only reported on a quarterly basis, therefore, the filings of a quarter are used for all days in the following quarter with a lag of 1 month to allow for a publication delay. Data corrections are implemented as in Daniel, Klos, and Rottke (2020a).

Table D.1: Relationship between streak portfolio returns and institutional ownership ratio

Each quarter, all stocks are sorted into terciles based on institutional ownership (IOR). Each day, we construct portfolios from the intersection with the five positive and five negative streak buckets. The table reports the value-weighted average portfolio excess-market returns, the Newey-West-adjusted t-statistics, the average number of stocks in the portfolio, and the number of sample days in which the portfolio was non-empty.

IOR		After negative streaks					After positive streaks				
Tercile		1	2	3	4	5	1	2	3	4	5
1 (low)	Ex. Mkt Ret. (in %)	0.021	0.078	0.114	0.17	0.194	-0.033	-0.033	-0.027	-0.043	-0.047
	t-stat	1.794	5.338	6.444	8.05	7.81	-2.159	-1.997	-1.415	-2.001	-1.64
	No. of stocks	733	354	167	79	37	671	292	126	55	24
	No. of days	5243	5243	5243	5243	5243	5243	5243	5243	5243	5242
2	Ex. Mkt Ret. (in %)	0.003	0.034	0.047	0.09	0.107	-0.009	-0.018	-0.024	-0.037	-0.062
	t-stat	0.45	3.825	3.897	5.747	5.582	-1.23	-1.88	-1.83	-2.527	-3.657
	No. of stocks	770	384	190	94	46	725	339	159	75	35
	No. of days	5243	5243	5243	5243	5243	5243	5243	5243	5243	5243
3 (high)	Ex. Mkt Ret. (in %)	0.017	0.052	0.077	0.107	0.132	-0.011	-0.045	-0.063	-0.062	-0.078
	t-stat	3.451	7.259	7.857	6.835	6.934	-2.387	-6.141	-6.79	-6.108	-5.259
	No. of stocks	783	395	198	100	50	751	363	175	84	41
	No. of days	5243	5243	5243	5243	5243	5243	5243	5243	5243	5243

[20] Our last IOR observation is from June 2018, slightly shortening the sample of portfolio returns to end in October 2018.

Table D.1 reports both the value-weighted streak portfolio returns sorted by the institutional ownership ratio (IOR). We observe higher portfolio returns among low IOR stocks than high IOR stocks. This result is broadly consistent with previous empirical results, showing that mispricing is concentrated in low IOR stocks. If one is willing to approximate the share of arbitrageurs in the economy with IOR, this result is also consistent with the model in Section 2, where mispricing due to return extrapolators is inversely proportional to the share of arbitrageurs in the economy.

Table D.2: Long-short streak strategy for high and low IOR

First, the universe of U.S. stocks is sorted based on the level of IOR. The sorted stocks are then separated into 3 buckets: the first bucket consists of all stocks with low IOR (below the 33% quantile), the second bucket with medium IOR level (above the 33% and below the 66% quantile), the third bucket with high IOR (above the 66% quantile). Within each of these buckets the long-short streak portfolios are computed. Reported are the time-series average of the market excess returns and the CAPM-, FF3-, and FF5-alphas. The final row reports the average difference in market excess returns or alphas of the streak strategy based on low IOR stocks minus high IOR stocks. Returns and alphas are reported in percentage. t-statistics are Newey-West-adjusted.

IOR-Tercile	Ex. Mkt Ret.	CAPM	FF3-alpha	FF5-alpha
Low IOR	0.177 (7.426)	0.168 (7.231)	0.170 (7.306)	0.178 (7.697)
2	0.105 (5.509)	0.097 (5.260)	0.098 (5.326)	0.111 (6.104)
High IOR	0.154 (10.120)	0.148 (9.992)	0.149 (10.118)	0.157 (10.751)
Low-High	0.023 (1.037)	0.020 (0.921)	0.020 (0.923)	0.021 (0.955)

We continue along the lines of Da, Huang, and Jin (2019) and test the relationship between streak portfolio returns and institutional ownership by testing the difference in performance of the long-short streak strategy based on stocks with an IOR level in the top third and an IOR level in the bottom third. The results are reported in Table D.2. Consistent with the results presented in Table D.1, the long-short portfolio returns are highest among low IOR stocks. However, the streak returns are smallest for a portfolio of medium IOR stocks and not for a portfolio of high IOR stocks.

The latter empirical observation is inconsistent with the theory laid out in Section 2, as long as IOR is a good proxy for the share of extrapolators in the population. It is also inconsistent with alternative explanations that interpret returns from streak strategies as a compensation of liquidity provision.