Does It Pay to Follow Anomalies Research? Machine Learning Approach with International Evidence

Ondrej Tobek[*] & Martin Hronec[†]

May 2018

ABSTRACT

We study out-of-sample returns on 153 anomalies in equities documented in academic literature. We show that individual anomalies originally identified in the US are also on average profitable in Europe, Japan, and Asia Pacific but their returns do not survive transaction costs. We show that a more sophisticated machine learning technique that aggregates all the anomalies into one mispricing signal is 4 times more profitable and survives on large cap universe of liquid stocks. The machine learning also leads to 2 times larger Sharpe ratios with respect to corresponding standard finance methods. It is thus profitable even after transaction costs in all the regions. We next study value of international evidence for selection of quantitative strategies that outperform out-of-sample. Past performance of quantitative strategies in the regions other than the US does not help to pick out-of-sample winning strategies in the US. Past evidence from the US, however, captures most of the predictability within the other regions. The value of international evidence in empirical asset pricing is thus very limited.

JEL classification: G11, G12, and G15.

[*]University of Cambridge, Department of Economics, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, UK, ot234@cam.ac.uk

[†]Charles University, Institute of Economic Studies, martin.hronec@fsv.cuni.cz

Low interest rates environment after Financial Crisis of 2008 has caused a surge in search for alternative ways how to earn steady returns that are uncorrelated with stocks market. One response of financial industry was an explosion in number of "smart beta" funds that provide exposure to various risk factors that have been historically connected to risk premia. This larger interest should, however, in turn lead to their lower profitability. McLean and Pontiff (2016) documented that post-publication returns on anomalies decrease by 58% relative to their in-sample returns. We thus evaluate profitability of portfolio level strategy that would attempt to invest in individual anomalies. This is a very realistic scenario as there are now many ETF funds replicating them. We focus on large capitalization universe of stocks where most of the quantitative funds operate and add crucial evidence by looking at the returns in regions outside the US. We then show that returns on anomalies can be significantly improved by synthesizing their information with machine learning techniques. We next turn to evaluation of role of international evidence for better selection of strategies that outperform out-of-sample. It is customary to assume that strategies that work everywhere should be of the highest interest to investors as they are very robust. We show that this does not have to be the case since these strategies should also attract the most attention which then drives their future returns down.

We first study portfolio level out-of-sample returns on a set of 153 published anomalies from which it is possible to construct long-short portfolios. We focus on both equal- and value-weighted returns on large cap universe of stocks. We restrict our universe of stocks to those with size larger than bottom decile in NYSE. Excluding micro-caps should lead to more relevant results for the quantitative strategies. We put larger emphasize on value-weighted returns because equal-weighted returns are susceptible to market microstructure biases as documented in Asparouhova, Bessembinder, and Kalcheva (2010). These biases can be substantial and can heavily influence the analysis. Green, Hand, and Zhang (2017) documented a significant drop in performance of all anomalies in the US after 2003. We observe a similar drop and a mixing strategy that equally invests in portfolios of all the significant published anomalies has insignificant positive value-weighted returns in the US since 1990. Anomalies identified in the US should be profitable outside the US if they approximate important systematic risks that command risk premia or if they capture behavioural biases that are general in nature. Examples of these are momentum of Jegadeesh and Titman (1993), book value to market value of Fama and French (1992), or maximum daily return in the past month of Bali, Cakici, and Whitelaw (2011). We confirm that this is indeed the case and the US market is perfect for empirical asset pricing studies due to its long sample.

We next shift our focus from the portfolio level analysis of individual anomalies to shrinkage methods that synthesize information from all anomalies into one mispricing signal. Gu, Kelly, and Xiu (2018) showed that machine learning methods can significantly outperform methods commonly employed in finance. We extend their finding from the US to international markets. Specifically, we consider a strategy that predicts future returns on individual stocks by their past characteristics (cross-sectional quantiles of anomalies). We estimate the regression on past data and predict next month returns from the latest characteristics. We then construct investment portfolios by buying top decile of stocks with the highest predicted return and short-selling stocks in the bottom decile predicted returns. We compare standard approach of using Fama and MacBeth (1973) least squares regressions, as in Lewellen et al. (2015), with gradient boosting regression trees. We find that machine learning can lead to significant gains. Sharpe ratio on value-weighted portfolio with gradient boosting is more than twice larger relative to least square method in the US.

We also evaluate usefulness of international evidence for estimation of post publication returns on the anomalies. Hou, Xue, and Zhang (2017) and Harvey, Liu, and Zhu (2016) showed that many anomalies cannot be replicated and many others are due to data mining. International evidence provides larger sample that should in turn lead to more powerful statistical tests. It also provides more relevant data which should be less susceptible to past structural changes in markets. One problem could be that some of the anomalies are specific to the US as they depend on the institutional setting. For example accruals depend on specific accounting rules. We show that the past international performance does not add anything in our portfolio level analysis after controlling for the past performance in the US. Problem with strategies that work everywhere is that they should also have larger decay out-of-sample as they attract more attention from investors. This is supported by the fact that the strategies that are significant in the US and Europe earn on average 47% lower returns out-of-sample relative to strategies that are significant only in the US. The past performance in both Europe and Japan helps to predict future performance in the respective regions. This supports the hypothesis that markets individual regions are institutionally specific which in turn makes some anomalies more powerful there. Machine learning supports this as there is no gain from estimating expected future returns in the US on past data outside the US with respect to focusing solely on the US. The forecast in other regions, however, gain from past data there.

We evaluate marginal value of new anomalies for out-of-sample predictions by comparing out-of-sample returns of the shrinkage strategy that synthesizes anomalies published before 1995, 2000, 2005, or 2010. Most of widely accepted risk factors have been published before 1995 such as those in Fama and French (1992) 3 factor model and momentum of Jegadeesh and Titman (1993). The new discoveries should thus have lower marginal explanatory power over time as all the low hanging fruit has been picked up. It is also possible that the vetting procedure that authors have to undergo during the publishing process limits these decreasing returns to new knowledge. We show that it is indeed the case and there is gradual increase in mean returns on the shrinkage strategy over 2010-2016 period with addition of the more recently published anomalies. Investors can thus benefit from following recent academic anomalies research. One possible explanation is that old widely accepted anomalies have gained a larger attention and their returns are thus depressed relative to the newer anomalies.

Limits to arbitrage could explain the profitability if it is not possible to invest into the mispriced stocks. We thus conduct several robustness checks. We decompose the returns of long-short portfolios into long and short legs. It is often impossible to short-sell due to insufficient supply of borrowable stocks. We, however, find that both the long and short legs of our shrinkage strategies posses investment opportunity with respect to returns on the market. Short-selling constraints cannot thus fully explain the profitability. We also further limit our large cap universe to stocks with monthly dollar trading volume over $100 million dollars. This should guarantee that the investments strategies have large capacity. We find that most of the profitability survives nonetheless. Finally, we study transaction costs on the investment strategies and conclude that it is not profitable to trade individual anomalies but the shrinkage strategies remain profitable.

Our focus is the closest to Jacobs and Müller (2017c) and Jacobs and Müller (2017a) who study international evidence for anomalies research. It is, however, different in many aspects. First, we focus on investable universe of stocks and put emphasis on liquidity which should make our results more relevant to any investor. Second, we investigate role of international evidence for future predictions of returns. Jacobs and Müller (2017c) and Jacobs and Müller (2017a) have focused on strategies that were utilizing only data in the respective regions. We also use more advanced machine learning techniques that significantly improve our out-of-sample predictions. Our study is the closest in methodology and application of machine learning techniques to Gu et al. (2018) who focused, however, only on the US.

Our contributions are multiple. First, we evaluate returns on individual anomalies after transaction costs. This is important since transaction costs are the easiest explanation for why the returns are not arbitraged away. We then provide international evidence that machine learning methods can improve forecasts of expected returns on individual stocks. This provides further evidence that finance academia can benefit from these methods. We also evaluate marginal value of recent anomalies with respect to the existing ones. We show that the new research is providing new information about cross-section of returns and it improves out-of-sample forecasts. Lastly, we evaluate the usefulness of international evidence against data mining in the US. The long sample in the US already contains most of quantitative information embedded in cross-section of stocks and international evidence is not adding anything to it. It is a natural thing to assume that strategies that work everywhere will outperform out-of-sample. We show that this is not generally true.

The paper is structured as follows. We start with methodology and data description in Section I. We study the profitability of investing in individual anomalies in various regions and look at whether outperforming strategies can be better selected with addition of past international performance in Section II. We then focus on the shrinkage strategies in Section III. The marginal value of new anomalies is studied in Section IV and role of limits of arbitrage in Section V. We conclude in Section VI.

I. Data and Methodology

A. Data

Our source of accounting and market data for the US is Merged CRSP/Compustat database from WRDS. The sample spans 1926 to 2016 period and contains all NYSE, Amex, and NASDAQ common stocks (CRSP share code 10 or 11). We adjust the returns for delisting following guidance in Hou et al. (2017).[1]

We take international data from Reuters Datastream. We filter data following Ince and Porter (2006), Lee (2011), and Griffin, Kelly, and Nardari (2010). The procedure comprises of manually checking names of the shares in the database for over 100 expressions describing their share class. We leave only primary quotes of ordinary shares of the companies with few exceptions where fundamental data in Datastream is linked to other share classes.[2] We also exclude all REITs. All the returns in this study are converted to US dollars. We delete daily returns for days when the stock market was closed in a given country. We further improve quality of data with procedures described in Tobek and Hronec (2018) and covered in Appendix A. Tobek and Hronec (2018) study implications of choice of fundamental database on measurements of performance for individual fundamental anomalies. They show that significance of anomalies varies over different data sources and this can completely change research inference. Studies of aggregated anomalies do not, however, suffer from these problems. Our further analysis is thus not impacted.

Our sample includes 30 countries, 22 are developed and 10 emerging. We sort the countries into 5 groups: Europe (E) - Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom; Japan (J); Asia Pacific (AP) - Australia, New Zealand, Hong Kong, and

[1]Specifically, we use return over the month if the delisting is on the last day of the month. Relevant delisting return is then added as a return over the next month. Then we use delisting return (DLRET) from monthly file if it is not missing. If it is missing then we use $(1 + ret_{cum}) * (1 + DLRET_d) - 1$, where ret_{cum} is cumulative return in the month of delisting and $DLRET_d$ is delisting return from the daily file. Lastly, we fill the gaps with $(1 + ret_{cum}) * (1 + DLRET_{avg}) - 1$, where $DLRET_{avg}$ is average delisting return for stocks with the same first digit of delisting code (DLSTCD). Hou et al. (2017) apply average over the past 5 years but we found this method very noisy and a single large outlier had huge impact on the average value.

[2]We closely follow the description in Griffin et al. (2010) on what shares are not common.

Singapore; Asia Pacific Emerging (AP EM) - India, Indonesia, Malaysia, Pakistan, Philippines, South Korea, Sri Lanka, Taiwan, and Thailand; and China (CN). Table I shows average, minimum, and maximum number of stocks in cross-section of the individual regions. Full sample category includes all the available stocks and large cap category only those with capitalization larger than the lowest decile in NYSE and price larger than $1 ($.1 for Asia Pacific and $.25 for emerging countries) at the end of the previous June.

Table I
Number of stocks in cross-section

	Full sample			Large cap		
	mean	min	max	mean	min	max
AP EM	7338	5860	8687	1089	571	1653
Asia Pacific	2430	1012	3706	551	321	896
China	1703	913	2788	1268	240	2728
Europe	5194	4440	6121	1976	1410	2945
Japan	3141	2074	3678	1541	1030	2313
USA	4768	1993	7525	2340	1234	3852

Another important source of data for our anomalies is I/B/E/S which we obtained from WRDS. We merge I/B/E/S on Datastream directly as it is one of Reuters databases and Datastream includes the respective tickers in its static file. The merger with CRSP is done indirectly through CUSIPs. We first try to merge on 8 digit CUSIP and then on 6 digit CUSIP if unsuccessful. We check the success of the merger manually by comparing tickers on exchanges and names of the companies. We transform all the variables in I/B/E/S to US dollars with original Reuters exchange rates which are provided by WRDS.

B. *Anomalies*

Our sample includes 153 anomalies published in academic studies. The full list is provided in Appendix B and their detailed description in online appendix. We select primarily anomalies that have been described in McLean and Pontiff (2016), Hou et al. (2017), or Harvey et al. (2016). We focus on anomalies that are valid in cross-section of stocks so that we can form portfolios out of them. We exclude anomalies that are specific to the US and which cannot be constructed outside the US.[3] Some anomalies also require the classification of industries such as Hou and Robinson (2006). The choice in the original papers is mostly with respect to SIC industry classification. We apply third level Datastream classification which sorts industries into 19 groups instead. This has one main reason. The industry classification in Datastream is available only from the static file which means that only the latest value is available. Variation over time for individual firms between closely related SIC codes would thus cause problems. We provide the transition between SIC classification and Datastream classification in the online appendix. There are 93 fundamental, 11 I/B/E/S, and 49 frictions anomalies in our sample. The anomalies come primarily from top 3 finance and top 3 accounting journals. Figure 1 graphs number of published anomalies over time. The second line is capturing number of anomalies whose in-sample period in their respective studies has ended. The number of anomalies has been gradually increasing over time without any apparent jumps.

[3]This includes anomalies: based on quarterly fundamental data since there is only short coverage internationally; connected to hand collected data in the US such as IPOs, SPOs, and mergers; requiring segment information and NBER data; and that are institutionally specific such as share turnover or effective tax rate. Some fundamental anomalies could not be implemented in Datastream as the required items are missing there.

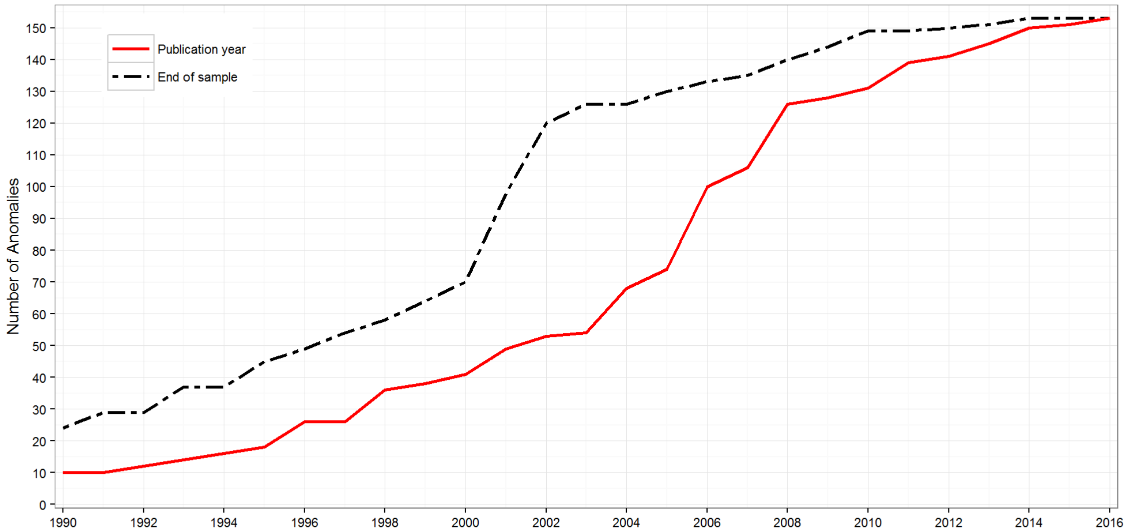


Figure 1. Number of published anomalies over time.

C. Portfolio construction

We primarily construct our portfolios at large cap universe of stocks if not stated otherwise. The focus on large cap universe should make the findings more realistic to someone trying to trade on the anomalies. The micro-caps account for a small fraction of the overall capitalization and often cannot be traded due to their high illiquidity. The second reason is that fundamental coverage of small cap stocks outside the US is very problematic and this could introduce huge biases into our analysis. We mostly provide both equal- and value-weighted returns. We prefer to focus on value-weighting. The main reason for this is that we don't want our analysis to be influenced by market microstructure biases as documented in Asparouhova et al. (2010).

Our portfolios on individual anomalies start in July 1963 in the US, 1990 in Europe, Japan, and Asia Pacific, and 2000 in China and Asia Pacific Emerging.[4] We omit the period before 1963 in the US for most of our analysis because the quality of returns and number of available stocks in CRSP is very low during that time. The coverage by Compustat is very low as well and this makes construction of most of the anomalies impossible there. We follow the original studies regarding further restrictions of the sample of stocks based on industries, age of firm, and number of years the firm's fundamental data have to available before it enters sample when construction portfolios on the individual anomalies. We also follow the original studies regarding rebalancing period so that most of I/B/E/S and frictions anomalies are rebalanced monthly, whereas, fundamental anomalies are rebalanced annually in July. Our zero-cost long-short portfolios on individual anomalies are constructed by buying stocks in the top decile of the signals and shorting stocks in the bottom decile of the signals.[5]

Portfolios that combine information from individual anomalies start in July 1995 unless stated otherwise. They are again long-short zero-cost but their returns now also correspond to a strategy that holds \$1 in cash, invests \$1 in the stocks that are likely have the largest return in the next

[4]International studies using fundamental data, such as Fama and French (2017), usually start in 1990. The reason for this is that there is insufficient fundamental coverage before that.

[5]We prefer zero-cost portfolios here since some annually rebalanced anomalies experience lower than -100% return during some years which creates problems with definition of return in terms of change in value of invested money with respect to previous month. We would have to introduce leverage constraints which would unnecessarily complicate of analysis.

month, and shorts $1 worth of stock that are likely to have the smallest return in the next month. The portfolios are then rebalanced to have equal share in cash, long, and short side of investment in stocks at the beginning of each month.

Table II
Average of time-series correlations of returns on portfolios created for the individual anomalies across the regions.

	USA	E	J	AP	CN	AP EM
USA	1.000	0.284	0.136	0.161	0.032	0.136
E	0.284	1.000	0.158	0.148	0.022	0.129
J	0.136	0.158	1.000	0.104	0.028	0.075
AP	0.161	0.148	0.104	1.000	0.066	0.121
CN	0.032	0.022	0.028	0.066	1.000	0.036
AP EM	0.136	0.129	0.075	0.121	0.036	1.000

We first look at correlation of returns on long-short portfolios created from the individual anomalies across different regions in Table II to better understand what to expect. It presents average of time-series correlations of identical anomalies across the regions. The anomalies in the US are the most correlated with anomalies in Europe at about 28.6%. The international evidence should be generally very useful as the anomalies are not closely related across the regions and can thus serve as independent source of information.

II. Portfolio level analysis of anomalies

A. *Profitability*

We start our study with portfolio level analysis of profitability of individual anomalies. This should give us a good sense of what to expect once we move to more complicated methods that synthesize information embedded in the individual anomalies to one signal. Table III presents returns on a portfolio mixing strategy that invests equally in all portfolios on anomalies that had significant returns with t-statistic larger than 1.96 in the US in the last June based on data available up to that point. That is, it corresponds to a setting where someone is following anomalies research, replicates the published findings, and equally invests into all published anomalies that he was able to replicate on large cap universe. We follow the performance of the mixed strategy in all our regions. Our out-of-sample forecasts begin in July 1990 for developed countries and July 2000 for emerging countries. Global strategy equally invests in the 4 developed regions. We adjust the performance of the portfolio mixing strategies by CAPM and Fama-French three and five factor models in the respective regions.

The table documents that the portfolio mixing strategy is not significant in the US for value-weighted returns. Its alpha with respect to the Fama-French five factor model is very close to zero. The profitability is much higher in the other regions and it is significant at 5% level there. This is despite the fact that the anomalies have been chosen in the US without any regard for evidence from the individual countries. The anomalies documented in literature are thus successful in capturing fundamental risks that yield risk premia.

We also focus on profitability in emerging markets. There are many reasons why the strategies should be more profitable there. Emerging countries are associated with more ideosyncratic risk as their economies are often dependent on fragile foreign investment flows. They are also susceptible to liquidity shocks such as during Asian Financial Crisis in late 1990s. Lower quality of data and

Table III
Out-of-sample performance of portfolio mixing strategies

The table shows returns of a strategy that equally invests in all the anomalies that are significant in the US at 5% level. We reselect the significant anomalies for the next year at the end of each June and consider only those that have been published by that time. The sample spans July 1990 (2000) to December 2016 for stocks in developed (emerging) and excludes stocks with capitalization smaller than the bottom decile of NYSE at the end of previous June. We adjust the performance of the mixing strategy for Fama-French factors from individual regions. Standard errors in t-statistics are HAC adjusted, as in Newey and West (1987) with 12 lags.

	July 1990+					July 2000+	
	USA	E	J	AP	Global	CN	AP EM
Panel A: Equal-weighted portfolios							
Mean Return	0.51	0.58	0.37	0.59	0.51	0.31	0.64
	(3.19)	(4.08)	(3.23)	(4.72)	(4.85)	(3.64)	(5.35)
CAPM alpha	0.63	0.61	0.37	0.62	0.56	0.30	0.66
	(3.74)	(4.26)	(3.62)	(5.10)	(5.56)	(3.69)	(6.22)
FF3 alpha	0.45	0.47	0.26	0.46	0.42	0.26	0.65
	(5.66)	(6.72)	(3.17)	(3.97)	(9.24)	(3.31)	(5.37)
FF5 alpha	0.27	0.27	0.27	0.32	0.34	0.30	0.62
	(3.21)	(3.43)	(3.46)	(2.98)	(6.28)	(3.49)	(5.56)
Panel B: Value-weighted portfolios							
Mean Return	0.14	0.44	0.41	0.30	0.32	0.20	0.29
	(0.93)	(3.30)	(2.53)	(1.24)	(3.69)	(2.31)	(2.14)
CAPM alpha	0.27	0.47	0.41	0.36	0.36	0.20	0.33
	(1.83)	(3.58)	(2.52)	(1.53)	(4.11)	(2.29)	(2.20)
FF3 alpha	0.18	0.38	0.45	0.36	0.30	0.16	0.34
	(1.63)	(3.35)	(2.65)	(1.53)	(3.61)	(1.80)	(2.34)
FF5 alpha	-0.01	0.23	0.46	0.10	0.22	0.24	0.33
	(-0.11)	(1.16)	(2.64)	(0.43)	(2.21)	(2.49)	(2.70)

harder enforcement of property rights can again increase riskiness. But is it really the case that the anomalies are more profitable there? Our limited evidence finds no support for this. The portfolio mixing strategy is the most profitable in Emerging countries in Asia Pacific for equal-weighted returns but there is no such outperformance in China.

Green et al. (2017) showed that the profitability of all anomalies has decreased significantly after 2003. We document the same thing in Figure 2 which presents evolution of cumulative returns on the portfolio mixing strategy since June 2002. The profitability of individual anomalies in the US has dropped to the point that they yielded only about 10% in this whole period. Other regions were more profitable. To conclude, it appears from portfolio level analysis that the US is the most efficient market. This can be due to its larger overall size and better data availability. Returns on anomalies in regions outside the US are positive and significant which suggests that it could be profitable to invest in them before transaction costs.

B. *Does international evidence help to pick out-of-sample winners?*

We have documented so far that the anomalies are profitable out-of-sample in all the regions if they were identified on the past data in the US. Can we use international data to better pick the winning strategies? The international markets' data should be useful against data mining since it provides larger sample so that all the statistical tests should have larger power. It also provides more relevant data coming from more recent time. This is important since the financial markets

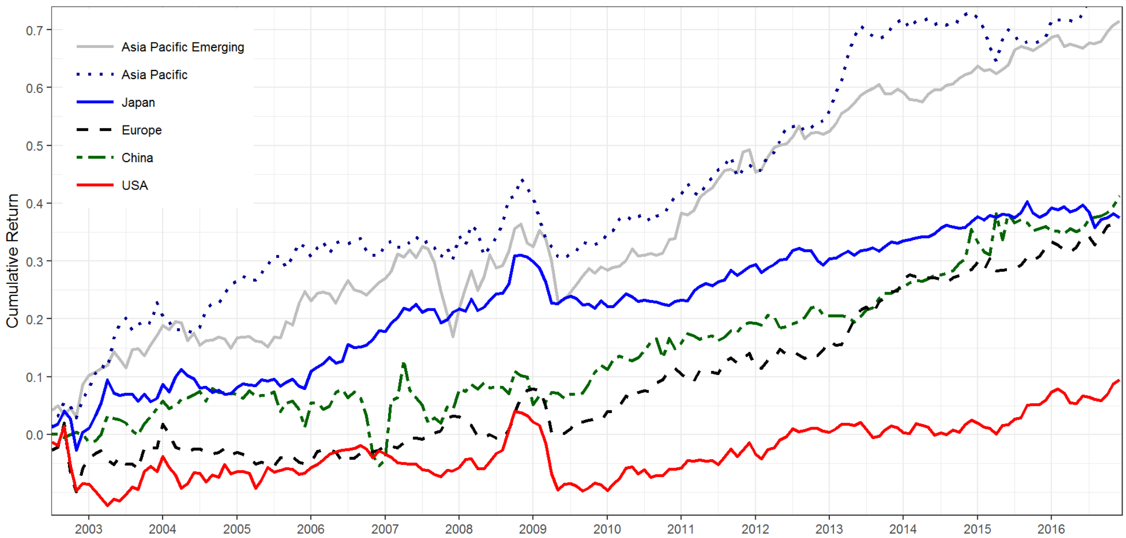


Figure 2. Performance of portfolio mixing strategy over different regions. The figure shows cumulative returns of a strategy that equally invests in all the published anomalies that are significant in the US at 5% level using the data available up to the end of prior month.

are changing rapidly and more relevant data should be more useful compared to evidence from 90 years ago. But there are also some downsides to its applicability. Different institutional specificities in the other regions can lower the usefulness in the US.

Table IV provides regressions of future t-statistics (or equivalently Sharpe ratios) on past t-statistics in Europe, USA, and Japan. We restrict our analysis to these three regions as there are only few large cap stocks in Asia Pacific historically. Panel A of the table uses simple OLS regression of t-statistics from the original sample of the publications regressed on post-publication five-year t-statistic in the three regions. We use post-publication t-statistic as returns are very noisy and do not lead to any significant finding. The methodology is suited to answer whether inclusion of international evidence in the original papers could have decreased the data-mining and post-publication decay. It is apparent that evidence from Japan and Europe is not useful for predictions of post publication performance in the US. The past performance in Europe is, however, somewhat useful for predictions in Europe and the same is also true for past performance in Japan. This supports the hypothesis that individual markets are institutionally unique to some degree. The predictability is much higher for equal-weighted portfolios relative value-weighting where none of the coefficients is significant at 5% level.

Panel B then studies the predictability in panel setting. We estimate past t-statistics in the three regions and future three-year t-statistics at the end of June every three years starting in 1995.[6] We fill all the missing t-statistics with zeros. Only anomalies that had been published at the time of selection are retained in the sample. All the regressions are then adjusted for time effects. The panel is suitable to answer whether international performance is useful in estimating future performance of published anomalies and thus to help to pick winning strategies at any point in time. The table again shows some predictability of future performance for their respective regions but only past performance in the US is useful when the past performance from all the regions is included together for value-weighted portfolios. Past evidence from Europe is, however, useful for equal-weighted portfolios in Europe. This puzzling result can be explained by the fact

[6] We have chosen to start our sample in 1995 so that there is enough evidence in Europe and Japan. The choice of starting point does not affect our conclusions.

Table IV
Predictive power of past performance in different regions

In Panel A, we estimate OLS regressions of 5-year post publication t-statistic on in-sample t-statistic in different regions for all the anomalies. We adjust the standard errors for heteroskedasticity. In Panel B, we regress future three-year t-statistics on past t-statistics for all the anomalies published up to that month. The t-statistics are using the longest sample available up to that date. The regression is based on panel data from the end of June 1995 to June 2014 at 3 years increments so that the out-of-sample periods are not overlapping. Standard errors in t-statistics are HAC adjusted, as in Newey and West (1987) with 12 lags. The sample spans July 1987 (1963) to December 2016 for international (US) stocks and excludes stocks with capitalization smaller than the bottom decile of NYSE at the end of previous June.

	Dependent variable from											
	USA				Europe				Japan			
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Panel A: 5-year post-publication t-statistics regressed on in-sample t-statistics												
Equal-weighted portfolios												
Intercept	0.14	0.57	0.45	0.15	0.64	1.26	1.38	0.76	0.12	0.56	0.64	0.07
	(0.73)	(3.83)	(3.98)	(0.78)	(2.61)	(6.22)	(7.98)	(2.74)	(0.62)	(3.01)	(4.34)	(0.33)
In-sample t-stat USA	0.19			0.27	0.29			0.32	0.24			0.31
	(2.71)			(3.85)	(3.90)			(3.41)	(3.39)			(3.95)
In-sample t-stat E		-0.09		-0.25		0.10		-0.08		0.09		-0.11
		(-1.40)		(-3.31)		(1.06)		(-0.88)		(1.04)		(-1.24)
In-sample t-stat J			0.05	0.07			0.11	0.07			0.19	0.15
			(0.64)	(1.06)			(0.96)	(0.64)			(2.31)	(1.89)
Sample Size	139	109	108	108	139	109	108	108	139	109	108	108
Value-weighted portfolios												
Intercept	0.36	0.56	0.41	0.53	0.43	0.76	0.72	0.62	0.22	0.40	0.45	0.14
	(1.78)	(3.02)	(3.34)	(2.14)	(2.24)	(4.45)	(5.57)	(2.68)	(1.28)	(2.53)	(3.50)	(0.69)
In-sample t-stat USA	0.02			0.04	0.16			0.14	0.16			0.20
	(0.28)			(0.42)	(1.63)			(1.38)	(1.81)			(1.89)
In-sample t-stat E		-0.19		-0.20		-0.07		-0.15		0.06		-0.02
		(-1.68)		(-1.83)		(-0.62)		(-1.28)		(0.52)		(-0.14)
In-sample t-stat J			-0.08	-0.01			0.07	0.11			0.03	0.02
			(-0.73)	(-0.11)			(0.47)	(0.73)			(0.33)	(0.18)
Sample Size	139	109	108	108	139	109	108	108	139	109	108	108
Panel B: Future 3-year t-statistics regressed on past t-statistics												
Equal-weighted portfolios												
Intercept	-0.03	0.24	0.36	-0.02	0.30	0.53	0.95	0.27	0.21	0.43	0.42	0.21
	(-0.23)	(4.77)	(22.30)	(-0.17)	(3.61)	(8.02)	(23.20)	(3.21)	(3.35)	(9.68)	(11.90)	(3.25)
In-sample t-stat USA	0.20			0.22	0.33			0.23	0.14			0.16
	(3.66)			(3.60)	(8.26)			(3.44)	(4.72)			(3.59)
In-sample t-stat E		0.10		-0.04		0.32		0.19		0.05		-0.07
		(2.70)		(-1.24)		(6.94)		(3.41)		(1.55)		(-1.30)
In-sample t-stat J			0.05	-0.01			0.06	-0.06			0.14	0.10
			(1.66)	(-0.36)			(0.81)	(-1.06)			(2.15)	(1.63)
Time Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Sample Size	532	532	532	532	532	532	532	532	529	529	529	529
Value-weighted portfolios												
Intercept	0.03	0.16	0.26	0.02	0.31	0.47	0.55	0.30	0.21	0.33	0.34	0.21
	(0.34)	(4.53)	(37)	(0.20)	(3.82)	(6.15)	(53)	(3.14)	(5.21)	(11.30)	(38)	(4.78)
In-sample t-stat USA	0.15			0.14	0.17			0.16	0.09			0.11
	(2.52)			(2.01)	(3.00)			(5.32)	(3.51)			(5.23)
In-sample t-stat E		0.11		0.06		0.10		0.03		0.02		-0.03
		(2.72)		(1.59)		(1.13)		(0.40)		(0.65)		(-0.84)
In-sample t-stat J			-0.02	-0.07			0.02	-0.02			0.01	-0.00
			(-0.82)	(-1.86)			(0.38)	(-0.64)			(0.36)	(-0.01)
Time Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Sample Size	532	532	532	532	532	532	532	532	529	529	529	529

that strategies that work everywhere should attract larger attention of investors who in turn drive their returns down. A simple test is to consider all the strategies that are significant in the US with t-statistic larger than 1.65. Anomalies that are also significant in the Europe provide 49% lower returns than those that are not.[7] One problem with this simple analysis is that anomalies

[7] There are 10 strategies that are significant in the US and Europe and 20 that are not. The sample is reduced due to missing in sample t-statistics in regions outside the US. There is only one strategy with t-statistic larger than 1.65 in all three regions, so the comparison there is not meaningful. The same is true for US and Japan with only 4 strategies.

identified later also tend to be more significant outside the US due to longer sample there.

Table V
Can international evidence improve out-of-sample performance of the portfolio mixing strategies?

The table shows returns of a mixing strategy that equally invests in individual portfolios of all significant anomalies using the data available up to the end of prior June and that were published by that time. The value-weighted or equal-weighted long-short portfolios on individual anomalies are constructed by buying stocks in the top decile of the signals and shorting stocks in the bottom decile of the signals. The significance is measured by t-statistic on mean returns on the portfolio in the US, US & Japan, US & Europe, and US & Japan & Europe. The sample spans July 1990 (1963) to December 2016 for international (US) stocks and excludes stocks with capitalization smaller than the bottom decile of NYSE at the end of previous June. The mixing strategy starts in July 1995. We adjust the performance of the mixing strategy for five Fama-French factors from individual regions. Standard errors in t-statistics are HAC adjusted, as in Newey and West (1987) with 12 lags. The returns are in percentage points per month.

	Equal-weighted				Value-weighted			
	USA	Europe	Japan	Asia Pacific	USA	Europe	Japan	Asia Pacific
Evidence from the US								
Mean Return	0.362	0.486	0.222	0.585	0.222	0.329	0.252	0.622
	(2.120)	(3.160)	(1.900)	(4.830)	(1.880)	(3.700)	(2.260)	(4.710)
FF5 alpha	0.172	0.214	0.127	0.315	0.098	0.070	0.275	0.461
	(1.870)	(2.800)	(2.110)	(2.900)	(1.100)	(0.536)	(2.140)	(2.960)
Evidence from the US & Japan								
Mean Return	0.357	0.470	0.247	0.615	0.234	0.323	0.304	0.605
	(2.140)	(3.210)	(2.290)	(5.540)	(1.640)	(3.560)	(2.390)	(4.560)
Diff wrt the US	-0.005	-0.016	0.024	0.030	0.011	-0.006	0.052	-0.017
	(-0.311)	(-0.853)	(0.866)	(0.798)	(0.216)	(-0.135)	(0.973)	(-0.351)
FF5 alpha	0.156	0.212	0.161	0.287	0.161	0.143	0.307	0.433
	(1.830)	(3.030)	(2.750)	(3.700)	(2.100)	(1.600)	(1.970)	(2.430)
Evidence from the US & Europe								
Mean Return	0.308	0.437	0.229	0.543	0.161	0.338	0.204	0.622
	(1.960)	(3.090)	(2.210)	(5.060)	(0.994)	(3.480)	(2.700)	(4.320)
Diff wrt the US	-0.054	-0.049	0.006	-0.042	-0.060	0.009	-0.049	-0.000
	(-1.480)	(-1.850)	(0.219)	(-1.010)	(-0.809)	(0.220)	(-0.553)	(-0.005)
FF5 alpha	0.125	0.213	0.137	0.248	-0.012	0.016	0.135	0.330
	(1.400)	(2.940)	(2.090)	(3.090)	(-0.155)	(0.134)	(1.710)	(2.780)
Evidence from the US & Japan & Europe								
Mean Return	0.343	0.459	0.254	0.544	0.191	0.297	0.263	0.616
	(2.180)	(3.280)	(2.800)	(5.620)	(1.010)	(2.980)	(2.880)	(4.260)
Diff wrt the US	-0.019	-0.027	0.032	-0.041	-0.031	-0.032	0.010	-0.006
	(-0.647)	(-0.787)	(0.592)	(-0.929)	(-0.308)	(-0.640)	(0.114)	(-0.090)
FF5 alpha	0.158	0.222	0.190	0.286	0.039	0.056	0.209	0.276
	(2.170)	(3.280)	(3.320)	(4.160)	(0.457)	(0.704)	(1.970)	(2.120)

Table V studies the benefit of past international evidence for better selection of strategies that should outperform out of sample. It shows mean returns and alphas with respect to factor models on portfolios created by equally combining individual portfolios for all the significant signals at 5% level for July 1995 to December 2016 period. The significance is determined by past t-statistics on intercept in panel regression of portfolio returns on just intercept. The portfolio returns are from the US, the US & Japan, the US & Europe, or the US & Japan & Europe. The t-statistics are HAC robust. The significant signals are chosen at the end of each June and the past t-statistics are taken for the period from July 1963 (1990) for the US (Japan and Europe) up to the time of portfolio formation. The table supports our evidence from the previous table in that the addition of international evidence does not lead to any significant improvement in out-of-sample performance of a strategy that mixes the most significant anomalies.

III. Shrinking anomalies into one mispricing signal

We have so far focused on portfolio level analysis of individual anomalies. we will now shift our attention to strategies that shrink all the anomalies into single mispricing signal ('shrinkage strategies' from now on). We follow Lewellen et al. (2015) in definition of the prediction problem. The goal is to devise a forecasting method that will predict which stocks are likely to have the highest returns in the next month and which the lowest. To do this, we regress past monthly individual stock returns on their characteristics available before the measurement period of returns. We normalize the characteristics to cross-sectional quantiles within each region. Normalization to quantiles should reduce problems with outliers and make the estimation more robust. We estimate these regressions by pooling all available stock returns up to date of portfolio formation. We then predict the future return from the latest available characteristics.

To summarize, we are estimating an equation

$$r_{it} = f(x_{i,t-1,1}, x_{i,t-1,2}, ..., x_{i,t-1,M}) + \epsilon_{it} \tag{1}$$

where r_{it} is returns on stock i in month t and $x_{i,t-1,1}$ is cross-sectional quantile of a given anomaly (characteristic) for the stock i available just before the start of month t. We demean the returns by subtracting average cross-sectional returns in every month and region. We start with a simpler case where $f()$ is linear and then extend it to a more general structure using machine learning.

A. Description of methods

A.1. Linear model

Our base model uses least square estimation for linear approximation of the relationship. That is we estimate weighted least square regressions

$$r_{it} = \beta_0 + \beta_1 x_{i,t-1,1} + \beta_2 x_{i,t-1,2} + ... + \beta_M x_{i,t-1,M} + \epsilon_{it} \tag{2}$$

where we weight by inverse of number of stocks in each time period and region. The weights are introduced to give the same role to each time period. This makes the moment conditions equivalent to Fama and MacBeth (1973) regressions in Lewellen et al. (2015) and should increase forecasting accuracy as most of uncertainty is hidden in time variation. The linear specification has already been applied in international context in Jacobs and Müller (2017c) and Jacobs and Müller (2017a). We thus use it as our benchmark for more complicated machine learning method.[8]

A.2. General model

We will now shift our focus from linear representation to the general case of nonlinear $f()$. Freyberger, Neuhierl, and Weber (2017) have proposed to use sparse additive models. This representation allows for each explanatory variable to have nonlinear effect but it does not allow for interaction effects without excessive computational burden. Gu et al. (2018) then applied a suite of standard machine learning algorithms and showed that they outperform both additive models and linear models in the US. We will focus on one of the machine learning methods here. We refer the readers to Gu et al. (2018) for detailed theoretical description of the machine learning method and we will here cover only basic description.

[8]We also tried volume-weighted regressions as in Green et al. (2017) which should put lower weight on small cap stocks and be more suited for value-weighted portfolio. They did not outperform simple regressions and we thus do not report the results.

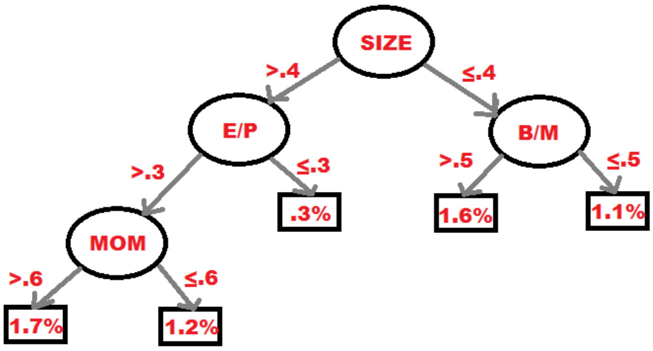


Figure 3. Decision tree.

We will focus on regression tree family of methods here as they are easy to estimate and require few specified parameters. One such tree is depicted in Figure 3. The decision tree consists of nodes (the round boxes) and outcomes (square boxes). The outcomes are in percent return per month but the numbers there are arbitrary and do not reflect real data. The tree starts with a decision whether a given stocks is withing the smallest 40% of stocks in cross-section. If it is the case then the decision continues based on book to market. The depicted tree is of depth 3, which is the maximum number of nodes in the longest branch. The tree allows for arbitrary cross-effects between the variables up to depth-1 degree. We will deal with relatively shallow trees here but they should still capture various important interactions between the explanatory variables. One large benefit of fitting simple trees is that they are easy to visualize.

The rest of the methods that we introduce are based on combination of the individual trees. These methods cannot be visualized but they lead to better out-of-sample forecasting performance. One of the most widely used ensemble method is a random forest. Random forest combines forecast from individual decision trees that are applied to a subsample of data. Explanatory variables are often also subsampled for the individual trees to increase variety among the individual forecasts. Random forest is frequently among top 10% of best performing machine learning methods in competitions and it is thus a very robust method that works most of the time. It also requires only very little specifications for parameters and can thus be used out of box. One downside is that it takes a long time to estimate. We thus focus on a simpler method with similar predictive ability.

Our analysis uses gradient boosting regression trees (GBRT) of Friedman (2001) which relies on a different way of combining the regression trees. All the trees in random forest are chosen independently whereas they are selected in dependent fashion in gradient boosting. The idea is to estimate a tree and use only a fraction fraction of its fit for forecasts. The next iterations then proceeds on residuals of dependent variables after removing the part of fitted values in the previous iteration. This should guarantee that the learning can correct itself if fitted values are selected incorrectly in some iterations. The fraction of individual predictions that is retained left is called a learning rate. Number of iterations given the learning rate then determines how closely we over-fit realization of the data in estimation sample. Selection of fewer iterations reduces the risk of over-fitting (estimation error) but decreases overall fit of the estimation (i.e. introduces an approximation error). It is thus important to select a reasonable number that will set a correct trade off. One way to do this, is to rely on cross-validation. The method thus requires a specification of learning rate, subsampling structure, number of iterations (trees), and maximum depth of the trees.

We do all our analysis with a quicker version of the gradient boosting - extreme gradient boosting of Chen and He (2017). The reason for this is that it is 10 times faster to estimate and thus does not require large computational power. It is also consistently among 10% of winning strategies in recent machine learning competitions. Gu et al. (2018) benchmarked the different machine learning methods and only neural networks provided better forecasts than GBRT. GBRT is thus a good candidate for the practical application and it captures most of the gain from machine learning over standard methods. We set the specification of regressions as follows. We choose shallow trees of maximum depth of 5 nodes. Gu et al. (2018) showed that cross-validation selects similar values in their analysis. Each iteration is done on a 50% subsample of the overall data. This should limit the over-fitting. We set the learning rate to 10% and run only 100 iterations.[9]

The machine learning methods have their benefits and negatives. They should provide better out-of-sample forecasts as they are built for it through limitation of in sample over-fitting. They also allow for very general interaction between the explanatory variables. This general form, however, makes the fitted models very hard to estimate and the methods thus remain black boxes. This should not be a large concern here since linear methods also become intractable given the number of exogenous variables and our metric is out-of-sample performance and not interpretation of estimated parameters. The machine learning methods also depend on specification of some parameters. A common approach in machine learning literature is to use cross-validation to choose then in data-dependent way. This is, however, not very successful in finance applications because the signal-to-noise ratio is tiny. Time dependence complicates it even more. We try to overcome these difficulties by setting constant parameters in gradient boosting in all our estimations. Cross-validation can thus only improve the reported returns if it posses any additional value.

A.3. Measure of out-of-sample predictability

We compare individual methods and forecasts based on data from different regions with out-of-sample (OOS) R^2 of the regressions. We follow Gu et al. (2018) and define absolute predictive ability as

$$1 - \frac{\sum_{it}(r_{it} - \hat{f}(x_{it}))^2}{\sum_{it} r_{it}^2} \tag{3}$$

where $\hat{f}$ is a predictive function fitted on data preceding month t. Unlike Gu et al. (2018), we use demeaned returns in each year instead of excess returns with respect to market return since we do the same in the regressions. We compare individual forecasts in the Diebold and Mariano (1995) test. Specifically, we adopt the approach in Gu et al. (2018) and create a time-series of differences in cross-sectional sums of the squared losses of the two forecasts. We then test significance of the differences by testing significance of their time-series mean with simple t-statistic. We adjust the standard errors in t-statistic for heteroskedasticity and autocorrelation with Newey-West procedure with 12 lags.

B. *Profitability of the strategies*

We now turn to empirical study of performance of the introduced the shrinkage methods. We expect that these methods would lead to higher returns in both absolute term and on risk adjusted basis relative to mixing of portfolios on individual anomalies, as was already shown in Jacobs and Müller (2017a) for least squares. The more sophisticated machine learning methods should then provide higher out-of-sample predictive capacity relative to least squares, as was documented in the US in Gu et al. (2018).

[9]This is again for the sake of speed and setting the learning rate lower with larger number of iterations could possibly further increase the power. There is, however, extreme amount of noise in financial data and the slower learning should thus not be necessary. We have tried to set lower rate of learning but failed to find any gain in performance ourselves.

Table VI
Performance of shrinkage strategies

This table shows out-of-sample performance of shrinkage strategies for signals from individual published anomalies. It is based on long-short decile portfolio from strategies that combines all the available signals through predictive least square or gradient boosting regressions of individual stocks returns on transformed characteristics. That is, we estimate pooled regressions of monthly stocks returns on cross-sectional quantiles of their characteristics observable before each month start and predict future returns from the latest available characteristics. The value-weighted or equal-weighted long-short portfolios are constructed by buying stocks in the top decile of the predicted next month returns and shorting stocks in the bottom decile of the predicted next month returns. We reestimate the regressions at the end of each June and consider only those anomalies that have been published by that time. The out-of-sample performance is observed in the US, Europe, Japan, and Asia Pacific. The sample spans July 1963 to December 2016 in the US and July 1990 to December 2016 in other regions. It excludes stocks with capitalization smaller than the bottom decile of NYSE or price lower than $1 ($.1 in Asia Pacific) at the end of previous June. Regressions in Panel A are conducted only on the past US data and future returns are predicted from them in all the regions, whereas, regressions in Panel B are done individually in each respective region. Reported returns are for July 1995 to December 2016 period. We adjust the performance of the shrinkage strategies for 5 Fama-French factors from individual regions. Standard errors in t-statistics are HAC adjusted, as in Newey and West (1987) with 12 lags.

	Equal-weighted				Value-weighted			
	USA	Europe	Japan	Asia Pacific	USA	Europe	Japan	Asia Pacific
Panel A: Estimated in the US								
WLS								
Mean Return	1.620	1.600	1.240	1.680	0.968	1.200	1.020	1.360
	(3.090)	(3.900)	(4.460)	(5.160)	(2.050)	(3.490)	(3.120)	(3.860)
FF5 alpha	1.060	0.875	1.020	1.400	0.477	0.571	0.823	0.941
	(4.080)	(3.960)	(4.040)	(4.320)	(1.490)	(2.420)	(2.540)	(2.540)
Gradient boosting regression trees								
Mean Return	2.260	2.060	1.710	2.230	2.020	1.390	1.010	2.150
	(4.560)	(5.470)	(6.360)	(6.000)	(4.590)	(4.390)	(3.300)	(4.990)
FF5 alpha	1.770	1.550	1.520	2.050	1.760	0.946	0.824	1.610
	(6.320)	(7.130)	(6.300)	(5.720)	(5.420)	(3.870)	(2.980)	(4.740)
Diff wrt OLS	0.636	0.460	0.465	0.553	1.050	0.197	-0.012	0.792
	(3.790)	(3.680)	(3.220)	(2.810)	(2.600)	(0.852)	(-0.034)	(2.220)
Panel B: Estimated in the individual regions								
WLS								
Mean Return	1.620	1.880	1.180	1.930	0.968	0.890	0.582	1.310
	(3.090)	(5.580)	(3.500)	(5.200)	(2.050)	(2.980)	(1.500)	(2.690)
FF5 alpha	1.060	1.070	1.030	1.450	0.477	0.082	0.445	0.797
	(4.080)	(3.420)	(3.440)	(3.230)	(1.490)	(0.191)	(1.280)	(1.520)
Gradient boosting regression trees								
Mean Return	2.260	2.190	1.860	2.040	2.020	1.320	1.200	1.930
	(4.560)	(6.920)	(6.360)	(5.920)	(4.590)	(5.180)	(3.760)	(4.540)
FF5 alpha	1.770	1.410	1.710	1.590	1.760	0.700	1.150	1.530
	(6.320)	(4.620)	(5.900)	(3.610)	(5.420)	(2.000)	(3.610)	(4.060)
Diff wrt OLS	0.636	0.309	0.680	0.112	1.050	0.429	0.623	0.615
	(3.790)	(2.690)	(4.130)	(0.446)	(2.600)	(2.680)	(2.160)	(1.420)

Table VI presents mean returns and alphas on portfolios created by the shrinkage strategies. We fit weighted least square and gradient boosting regressions on returns and characteristics available up to June every year and then predict returns in the next month with the latest available characteristics for each of the next 12 months. We estimate the regressions on data from the US in Panel A and with data from the respective regions in Panel B. The estimates are based on returns from July 1963 (1990) to June 2016 in the US (elsewhere). We then create long-short decile port-

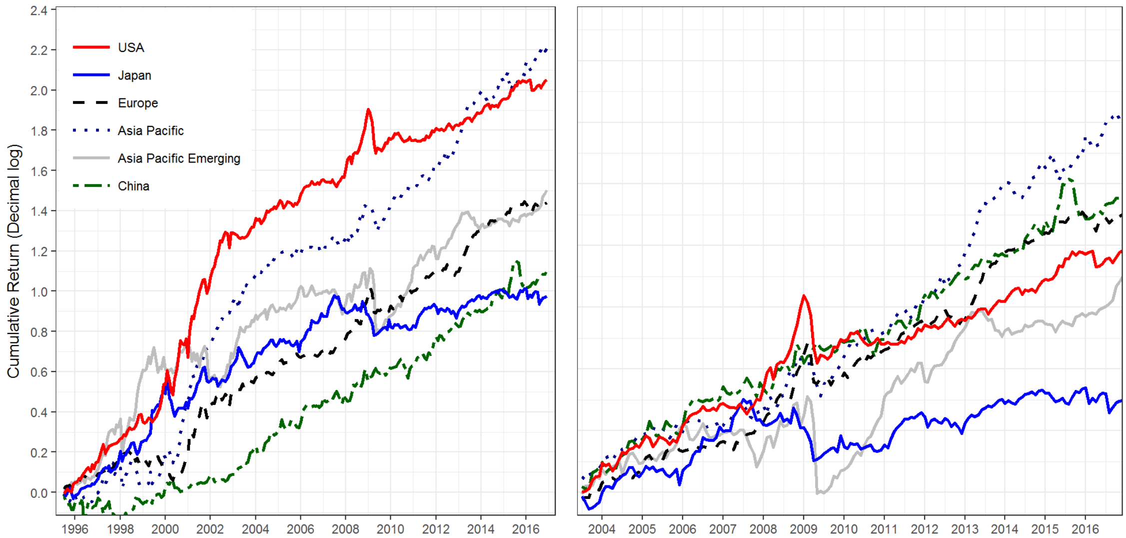


Figure 4. Cumulative returns on gradient boosting shrinkage strategy. The figure shows cumulative returns for the shrinkage strategy as described in Table VI that is estimated on individual stocks from the US.

folios that invest into stocks in the top decile of predicted future returns and short stocks in the bottom decile of predicted returns. The reported returns on portfolios are in percents per month and are from July 1995 to December 2016.

Gradient boosting outperforms least squares almost everywhere for both mean returns and risk adjusted Sharpe ratios. The more complicated machine learning method is thus better for predictions outside the US as well as inside the US. This provides robustness to finding in Gu et al. (2018). The average returns on the shrinkage strategies are about 4 times higher than for portfolio level analysis in the previous section. This is in line with Jacobs and Müller (2017a). Gradient boosting also outperforms WLS based on alphas with respect to 5 Fama-French factors. The nonlinear method thus better captures drivers of returns that are orthogonal to standard risk factors.

There is surprisingly only small difference between returns on strategies that are estimated on data from the US in Panel A and those that are estimated on data in the respective regions in Panel B. One explanation for this can be that the sample size in the US is already large enough to capture true drivers of stock returns that are globally valid. It is none the less surprising given that one would expect large heterogeneity in performance of anomalies in different regions due to institutional specificities.

Figure 4 plots cumulative returns on the gradient boosting strategy from Panel A in Table VI. The returns are in decimal logarithms and 1 thus corresponds to 10 times returns on initial investment. There is no apparent drop in profitability around 2003 in the US as for individual strategies in Figure 2. Strategies in the US and Asia Pacific have been historically the most profitable but there is only small difference among the regions after 2003. It is also apparent that the emerging markets are not significantly more profitable than the other regions.

C. Does international evidence help to pick out-of-sample winners?

The portfolio level analysis in the previous section found little value for international evidence on individual anomalies. We revisit the question here with the machine learning methods. The previous machine learning evidence was based on predictions with relationships estimated in the US or in the respective regions. We will now investigate whether pooling international and US stocks in the estimation can improve the predictions.

C.1. Returns on shrinkage strategies estimated on international data

Panel A in table VII shows mean returns and alphas on extreme gradient boosting shrinkage method as in Panel A in Table VI. The only difference is that the future returns are predicted from regressions estimated on stocks that are not only from the US. Specifically, we compare the cases with estimation done in the US, US & Japan, US & Europe, or US & Japan & Europe& Asia Pacific. This covers most of the developed markets and global capitalization. Corresponding evidence for least square shrinkage is provided in the online appendix.[10]

Why should we expect that the international evidence is useful? Standard statistical theory predicts that larger sample should allow closer convergence to the truth if our methods are consistent. One problem could be that the sample in the US is already large enough and we thus already know the truth from the US alone. This can easily be true as there is one million observations there. Most of the information for predictability of stocks, however, depends on sample length and not on number of stocks in the cross-section. This is because we can well explain realized returns within individual months but these fitted values vary greatly over time. The sample length should thus not be great enough to contain all the necessary information. It is also possible that international evidence will not prove to be useful since anomalies can be specific to individual regions given large institutional differences among the regions. An example is accounting based anomalies that depend on specific accounting standards. The institutional differences should hopefully not be too important as the US evidence would thus have lower predictive power relative to evidence from the individual regions when forecasting future returns in the respective regions. We previously showed that this is not the case.

The table provides mixed results on the value international evidence. There is no gain for predictions in the US. Historical data in the US is thus sufficient for the future predictions. Europe gains from predictions based on the data from the US and Europe relative to the US only. The same is also true for Japan and data from the US and Japan. The largest gains are in Asia Pacific that appears to benefit from data from both Japan and Europe. We thus find that the regional specificities are indeed important drivers but there is no gain for the US investor to seek international evidence for quantitative strategies.

C.2. Returns predictability for all the stocks

We have documented that international evidence has a limited value for profitability of long-short strategies. Will now examine whether this is true only in the extreme deciles of predictions or if it applies to forecasts of future returns on all stocks. To do this, we compare individual forecasts by their out-of-sample R^2. We follow our previous analysis in that we reestimate our regressions every year in June and then predict future returns for the next 12 months, one month at a time.

Panel B in table VII presents the results. It is immediately apparent that GBRT outperform WLS. This is significant with t-statistic larger than 4 and the outperformance remains even when stocks from all the regions are used in the estimation of the shrinkage strategies. WLS gains from the international evidence and the OOS R^2 significantly improve with large sample. The conclusion is not so clear for GBRT. There is no gain from international evidence in the US, as previously for returns on portfolios, but the other regions gain from the international evidence. Evidence

[10] We omit it here for the sake of space as all the findings are very similar to gradient boosting.

Table VII
Performance of shrinkage strategies estimated on stocks outside the US

The table shows returns (Panel A) and OOS R^2 (Panel B) of the shrinkage strategies described in Table VI that are estimated on individual stocks from the US, Japan, Europe, Europe, or Asia Pacific and their combinations. We adjust the performance of the mixing strategy for five Fama-French factors from individual regions. Standard errors in t-statistics are HAC adjusted, as in Newey and West (1987) with 12 lags. The returns are in percentage points per month. The OOS R^2 are in percentage points.

Panel A: Returns on the portfolios estimated with Gradient boosting regression trees

	Equal-weighted				Value-weighted			
	USA	Europe	Japan	Asia Pacific	USA	Europe	Japan	Asia Pacific
Estimated in the US								
Mean Return	2.260	2.060	1.710	2.230	2.020	1.390	1.010	2.150
	(5.820)	(6.730)	(6.840)	(7.010)	(4.960)	(4.760)	(3.090)	(5.460)
FF5 alpha	1.770	1.550	1.520	2.050	1.760	0.946	0.824	1.610
	(6.210)	(7.240)	(6.760)	(6.970)	(5.140)	(3.780)	(2.820)	(4.610)
Estimated in the US & Japan								
Mean Return	2.160	1.930	2.000	2.070	1.530	1.520	1.600	2.500
	(5.790)	(6.750)	(7.920)	(6.640)	(4.020)	(5.340)	(4.280)	(5.650)
Diff wrt the US	-0.096	-0.122	0.291	-0.162	-0.495	0.132	0.593	0.342
	(-1.000)	(-0.965)	(2.230)	(-0.985)	(-2.570)	(0.773)	(2.330)	(0.874)
FF5 alpha	1.650	1.460	1.860	1.770	1.120	1.040	1.550	2.170
	(6.120)	(7.100)	(7.920)	(5.710)	(3.110)	(3.970)	(4.170)	(4.200)
Estimated in the US & Europe								
Mean Return	2.170	2.340	1.740	2.590	1.930	1.460	1.170	2.710
	(5.450)	(7.190)	(6.490)	(7.270)	(4.590)	(4.710)	(3.130)	(6.250)
Diff wrt the US	-0.092	0.289	0.033	0.361	-0.094	0.062	0.163	0.554
	(-1.150)	(3.090)	(0.399)	(2.320)	(-0.436)	(0.376)	(0.731)	(2.180)
FF5 alpha	1.690	1.690	1.560	2.290	1.820	0.982	1.050	2.320
	(5.190)	(6.880)	(6.410)	(6.470)	(4.230)	(3.420)	(3.140)	(5.060)
Estimated in the US & Japan & Europe & Asia Pacific								
Mean Return	2.110	2.220	1.810	2.810	1.770	1.480	1.380	3.390
	(5.070)	(7.160)	(7.160)	(8.360)	(3.730)	(4.420)	(3.230)	(7.470)
Diff wrt the US	-0.150	0.164	0.099	0.584	-0.249	0.092	0.368	1.240
	(-1.450)	(1.560)	(1.030)	(3.450)	(-0.936)	(0.435)	(1.380)	(3.680)
FF5 alpha	1.600	1.530	1.670	2.520	1.440	0.945	1.420	3.160
	(4.890)	(7.250)	(7.100)	(7.110)	(3.610)	(3.230)	(3.560)	(6.750)

Panel B: Out-of-sample R^2

	WLS				Gradient boosting regression trees			
	USA	Europe	Japan	Asia Pacific	USA	Europe	Japan	Asia Pacific
Estimated in the US								
OOS R^2	0.026	0.045	-0.111	0.137	0.204	0.359	0.182	0.315
Estimated in Europe								
OOS R^2	-0.114	0.243	-0.212	0.486	-0.056	0.349	-0.074	0.570
Diff wrt the US	-1.650	2.240	-1.010	4.650	-3.570	-0.128	-2.820	3.490
Estimated in Japan								
OOS R^2	-0.763	-1.510	-0.500	-1.430	-0.709	-1.280	-0.104	-1.350
Diff wrt the US	-6.640	-9.260	-1.970	-8.180	-6.620	-8.780	-1.570	-7.880
Estimated in Asia Pacific								
OOS R^2	-0.327	-0.081	-0.674	0.355	-0.298	-0.214	-0.563	0.248
Diff wrt the US	-2.770	-1.120	-4.190	2.270	-5.640	-5.180	-6.650	-0.756
Estimated in the US & Europe								
OOS R^2	0.065	0.236	-0.020	0.314	0.199	0.446	0.219	0.439
Diff wrt the US	2.280	6.560	3.130	7.140	-0.309	3.030	1.110	4.600
Estimated in the US & Japan								
OOS R^2	-0.004	-0.065	0.027	0.007	0.167	0.222	0.324	0.199
Diff wrt the US	-1.270	-3.210	3.090	-4.180	-1.650	-3.380	3.200	-3.770
Estimated in the US & Japan & Europe & Asia Pacific								
OOS R^2	0.073	0.195	0.054	0.295	0.190	0.396	0.272	0.437
Diff wrt the US	2.050	4.220	3.890	4.410	-0.585	0.916	2.200	3.200

from the individual regions in not sufficient due to shorter samples and combining it with the past US data always improves the R^2. Our conclusions from the portfolio level analysis are thus very similar to OOS R^2 on all the stocks.

IV. Are the recent anomalies important?

Figure 1 has documented that the number of anomalies is increasing roughly linearly over time. Harvey et al. (2016) found even sharper increase for both published and unpublished anomalies that was closer to an exponential function. Researchers are looking at the same data again and again which should lead to a large proportion of false positive discoveries that is increasing over time as all the low hanging fruit is already gone. They have thus concluded that most of recently published studies can probably be explained as pure data mining. This should than lead to a lower predictive power of new anomalies. Many of the new anomalies are also subsumed by existing anomalies in proper multiple hypothesis setting. See, for example, Green et al. (2017). Detailed scrutiny during the publishing process should, however, limit these data mining issues. Most of the widely accepted anomalies, such as those in Fama and French (1992) three factor model, have been published before 1995. It is thus worth studying if the less widely accepted drivers of returns are also important. We will now investigate marginal value of recently published anomalies for profitability of the machine learning strategies after accounting for anomalies published earlier.

Table VIII presents mean returns on machine learning strategy applied in the Panel A in Table VI with further restrictions on universe of anomalies. We study out-of-sample performance for the shrinkage strategy based on anomalies published before 1995, 2000, 2005, and 2010 over 2005-2016 and 2010-2016 period. This should give us good indication of how valuable the new signals are if we account for the existing signals. It is apparent that there are a improvements in mean returns on both equal- and value-weighted portfolios in the US as we add more recent anomalies. The new anomalies thus have significant incremental value for out-of-sample forecasts. This benefit is much smaller in Japan and Europe. Asia Pacific, however, closely follows the US and there are strong benefits from recent anomalies there. The results are very similar for both least squares and gradient boosting methods. One explanation for the larger value in the US with respect to Europe and Japan is that there are more existing "smart beta" ETFs in the US that arbitrage away the well known strategies. It is thus necessary to find new strategies to get the same predictability. We test this in the next section while studying post-publication decay on anomalies.

To conclude, there is no apparent drop in marginal value of adding the new anomalies over time. Following recent academic research thus has value and increases possible returns to investors. This is in line with the purpouse of academic publishing process where new findings are put under deep scrutiny so that the authors have to prove that their findings provide incremental value with respect to existing evidence.

V. Limits to arbitrage

This section studies various possible explanations of the profitability of anomalies. We begin with decomposition of returns on long-short portfolios to long and short legs. This should allow us to see if most of the profitability is on the short side where it could be explained by shorting costs and sometimes impossibility to short-sell at all. We next study various restrictions on universe of stocks that should improve capacity of the strategies and decrease transaction costs on the strategies. Finally, we turn to post-publication decay as documented in McLean and Pontiff (2016)

Table VIII
Is marginal return to following new anomalies decreasing over time?

The table shows returns of the shrinkage strategies described in Table VI that are estimated on the individual stocks from the US. We restrict anomalies in the estimation to those that were published before 1995, 2000, 2005, or 2010. Standard errors in t-statistics are HAC adjusted, as in Newey and West (1987) with 12 lags. The returns are in percentage points per month.

	Equal-weighted				Value-weighted			
	USA	Europe	Japan	Asia Pacific	USA	Europe	Japan	Asia Pacific
Panel A: Mean returns over 2005-2016								
WLS								
Published before 1995	0.177	1.220	0.531	0.639	0.132	0.829	0.938	0.557
	(0.409)	(5.670)	(1.660)	(1.190)	(0.317)	(2.790)	(2.780)	(0.975)
Published before 2000	0.431	1.150	0.482	0.962	0.268	0.899	0.841	0.746
	(1.050)	(5.330)	(1.320)	(2.100)	(0.823)	(3.010)	(3.450)	(1.970)
Published before 2005	0.869	1.370	0.722	1.130	0.633	0.701	1.070	0.840
	(2.680)	(7.350)	(2.120)	(2.650)	(2.900)	(2.230)	(4.170)	(2.750)
Gradient boosting regression trees								
Published before 1995	0.474	1.770	0.884	0.985	0.391	1.050	1.350	0.942
	(1.400)	(6.880)	(2.700)	(1.910)	(0.882)	(2.830)	(4.130)	(2.120)
Published before 2000	0.741	1.680	0.876	1.450	0.509	0.968	0.974	1.380
	(1.950)	(6.870)	(2.500)	(2.960)	(1.300)	(3.260)	(2.760)	(3.430)
Published before 2005	0.987	1.860	0.946	2.000	1.160	0.970	1.170	1.710
	(3.270)	(9.860)	(3.270)	(3.630)	(3.530)	(2.380)	(3.210)	(3.440)
Panel B: Mean returns over 2010-2016								
WLS								
Published before 1995	0.431	1.370	0.455	1.150	0.269	0.955	1.000	0.692
	(2.140)	(4.970)	(1.460)	(2.410)	(1.060)	(2.640)	(2.470)	(2.020)
Published before 2000	0.667	1.280	0.541	1.530	0.206	1.050	0.929	0.757
	(2.510)	(4.140)	(1.590)	(3.660)	(0.777)	(2.630)	(2.920)	(1.900)
Published before 2005	1.090	1.620	0.596	1.740	0.951	1.090	1.040	0.768
	(5.240)	(6.650)	(2.050)	(4.570)	(4.330)	(2.340)	(2.990)	(1.720)
Published before 2010	1.340	1.910	0.888	2.520	0.804	1.230	0.924	1.630
	(5.980)	(7.530)	(4.290)	(4.490)	(3.710)	(2.330)	(3.260)	(5.200)
Gradient boosting regression trees								
Published before 1995	0.675	1.840	0.958	1.740	0.213	1.280	1.600	1.330
	(3.120)	(5.850)	(4.620)	(3.600)	(0.737)	(3.270)	(5.010)	(2.640)
Published before 2000	1.070	1.700	0.989	2.160	0.473	1.210	1.270	1.540
	(4.340)	(6.000)	(4.660)	(5.670)	(1.410)	(4.120)	(3.650)	(3.180)
Published before 2005	1.280	1.840	0.984	2.860	1.090	1.380	1.270	2.340
	(5.640)	(7.950)	(5.060)	(6.060)	(4.420)	(4.090)	(3.030)	(4.230)
Published before 2010	1.690	2.220	0.975	2.940	1.080	1.630	0.874	2.470
	(6.940)	(9.710)	(4.890)	(6.580)	(5.660)	(4.320)	(3.070)	(5.570)

on our universe of stocks and anomalies. We then extend it to decay on our shrinkage strategies.

A. Profitability of long and short sides of the investment strategies

Short-selling is often impossible or connected to large costs. It is also frequently outside mandates of asset managers. We here decompose the long-short strategy in Table VI to see if the high expected returns are due to short side or if it is possible to gain from the information embedded in the anomalies for long-only investments. The table decomposes the long-short returns separately for least squares shrinkage strategy and gradient boosting strategy. We also add equal- and value-weighted returns on the whole market as defined by our large cap universe of stocks. We provide both equal-weighted and value-weighted returns in each region in the respective columns in the table.

It is apparent that the gains from the strategy are valid for both long and short side. The returns on long leg of extreme gradient boosting strategy are about 10% a year larger than on the

Table IX
Decomposition of performance of shrinkage strategies to long and short legs

The table shows mean returns of the shrinkage strategies described in Table VI that are estimated on the individual stocks from the US. We decompose returns on the long-short portfolios to long and short legs. We also provide either equal- or value-weighted mean returns on the whole markets in individual regions that are estimated on our sample of large cap stocks. Standard errors in t-statistics are HAC adjusted, as in Newey and West (1987) with 12 lags. The returns are in percentage points per month.

	Equal-weighted				Value-weighted			
	USA	Europe	Japan	Asia Pacific	USA	Europe	Japan	Asia Pacific
				WLS				
Long leg	1.570	1.410	0.905	1.500	1.140	1.240	0.739	1.530
	(4.830)	(3.410)	(2.220)	(3.130)	(3.770)	(3.210)	(1.840)	(3.270)
Short leg	-0.056	-0.186	-0.340	-0.174	0.173	0.043	-0.284	0.168
	(-0.094)	(-0.318)	(-0.658)	(-0.271)	(0.316)	(0.092)	(-0.550)	(0.291)
				Gradient boosting regression trees				
Long leg	1.810	1.570	1.150	1.650	1.670	1.280	0.741	1.580
	(4.470)	(3.650)	(2.650)	(3.240)	(4.440)	(2.960)	(1.460)	(2.810)
Short leg	-0.451	-0.490	-0.556	-0.575	-0.353	-0.113	-0.270	-0.570
	(-0.727)	(-0.762)	(-1.040)	(-0.777)	(-0.566)	(-0.193)	(-0.574)	(-0.744)
				Whole market				
Long leg	0.987	0.798	0.365	0.802	0.825	0.708	0.247	0.834
	(2.900)	(1.900)	(0.930)	(1.600)	(2.660)	(1.930)	(0.616)	(2.040)

market. The same is also true for the least squares shrinkage strategy to a smaller extent. The extreme gradient boosting outperforms least squares shrinkage strategy primarily on the short side where it is able to identify stocks with negative returns. To conclude, the positive returns on our shrinkage strategies are robust to short-selling constrains and are available for all investors.

B. *Restrictions on universe of stocks*

All of our analysis so far has focused on large cap universe of stocks. We will here introduce further restrictions on the universe that will guarantee that it is possible to invest large amount of capital into the shrinkage strategies and we will study how it affects the returns. Table X introduces two new restrictions on the universe. First, we exclude stocks with price lower than $5 dollars at the time of portfolio formation. Stocks with low price are often hard or even impossible to short-sell and stock brokers often put strict leverage constraints on them. Pástor and Stambaugh (2003) are an example of a study that has introduced the same constraint. The second constraint discards all stocks with dollar trading volume lower than $100 million. This is to guarantee that the investment strategies have large capacity. Only the most liquid stocks with tiny transaction costs should satisfy this criterion.

Panel A documents that the average monthly returns on the shrinkage strategies are much lower for stocks with large trading volume. They nevertheless remain positive and mostly also significant. Panel B then shows that the drop in average returns is even smaller when stocks with price under $5 are excluded. This is due to the fact that we have already excluded stocks with price under $1 in our previous analysis, which has far larger impact. One notable difference is for strategies in Asia Pacific, as most of the stocks there are with only very small prices, and this restriction limits the universe of stock to only about 50 for some years. Combining the two restrictions in Panel C then documents that the significantly positive returns survive. Our previous finding are thus very robust and apply even to very liquid sample of stocks.

Table X
Performance of shrinkage strategies on restricted sample

The table shows mean returns of returns of the shrinkage strategies described in Table VI that are estimated on the individual stocks from the US. We restrict our large cap universe of stocks to those with price larger than $5 at the end the last month or/and with trading dollar volume of at least $100 million. Standard errors in t-statistics are HAC adjusted, as in Newey and West (1987) with 12 lags. The returns are in percentage points per month.

	Equal-weighted				Value-weighted			
	USA	Europe	Japan	Asia Pacific	USA	Europe	Japan	Asia Pacific
Panel A: Volume in the previous month larger than $100 million								
Mean Return WLS	1.020	0.995	0.480	0.683	0.760	0.750	0.517	0.476
	(2.170)	(2.750)	(1.080)	(1.810)	(1.630)	(2.460)	(1.350)	(1.320)
Mean Return GBRT	1.430	1.510	1.200	1.380	1.640	1.090	0.606	1.400
	(4.350)	(4.570)	(3.430)	(2.860)	(4.770)	(3.810)	(1.470)	(2.780)
Panel B: Price larger than $5								
Mean Return WLS	1.570	1.560	1.200	1.030	0.920	1.090	0.854	0.829
	(3.290)	(3.840)	(4.430)	(2.900)	(2.130)	(3.210)	(2.560)	(1.890)
Mean Return GBRT	1.980	2.030	1.720	1.310	1.890	1.200	1.040	0.944
	(4.530)	(6.040)	(6.100)	(4.010)	(4.910)	(4.470)	(2.980)	(2.080)
Panel C: Price larger than $5 and volume in the previous month larger than $100 million								
Mean Return WLS	1.030	0.840	0.661	1.000	0.761	0.751	0.669	1.060
	(2.290)	(2.200)	(1.680)	(2.760)	(1.660)	(2.470)	(1.760)	(2.660)
Mean Return GBRT	1.440	1.220	1.330	0.852	1.610	0.823	0.724	0.807
	(4.320)	(4.670)	(3.680)	(1.890)	(4.900)	(3.400)	(1.660)	(1.460)

C. Transaction costs on the strategies

We next study performance of out-of-sample strategies after transaction costs. Is could be the case that the profits on the strategies are only virtual and transaction costs are larger than the returns. We start with portfolio level analysis of our simple strategy that mixes individual anomalies and continue with our shrinkage strategies.

C.1. Portfolio mixing strategies

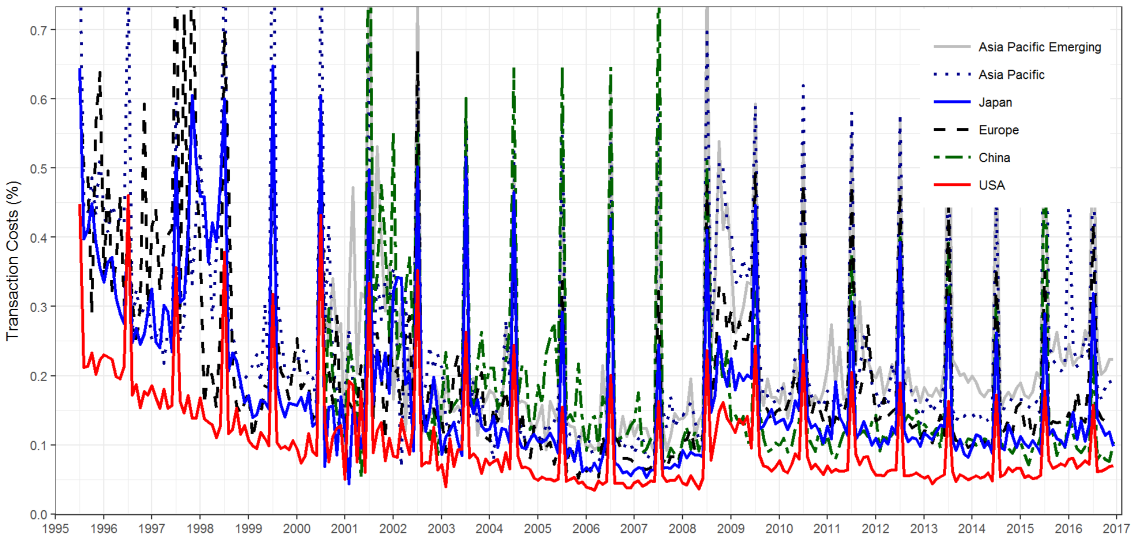
We have discussed in previous analysis that individual anomalies are more profitable outside the US. Does this superior performance persist after transaction costs? There are several reasons why the transaction costs should be higher outside the US. One of them is that the stocks in the US together constitute much larger capitalization this leads to larger average size of stocks in portfolios.

Panel A in Figure 5 describes transaction costs on the mixing strategy introduced in Section II. We measure transaction costs by VoV(% Spread) proxy introduced in Fong et al. (2017). It is defined as

$$8\,\frac{\sigma^{2/3}}{avg\ vol^{1/3}} \tag{4}$$

where σ is standard deviation of daily returns and *avgvol* average daily trading volume in USD within a given month. The trading volume is in USD and deflated to 2000 prices. The proxy roughly measures fixed component of trading costs and excludes price impact which would further increase the transaction costs. It roughly corresponds to estimated relationship between transaction costs and volume on large institutional portfolio transfers over as estimated by Kyle and Obizhaeva (2016) over 2002-2005 period. Fong et al. (2017) show that price impact component is very hard to measure, volatile over regions, and thus is very dependent on execution strategy of individual

Panel A: Mixing portfolios on individual anomalies.



Panel B: Gradient boosting shrinkage strategy.

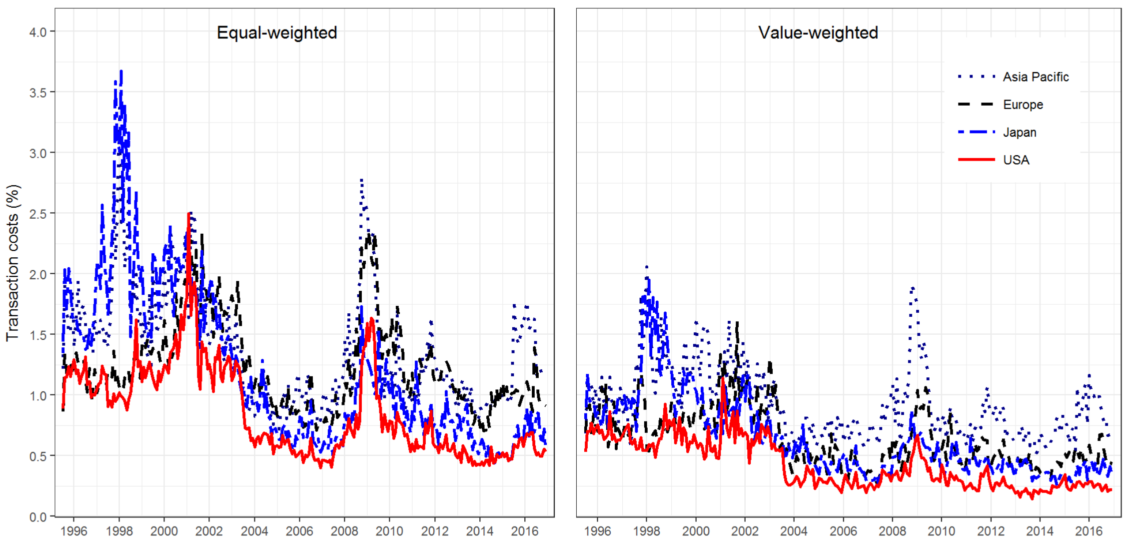


Figure 5. Average monthly transaction costs over different regions. Panel A shows transaction costs for the portfolio mixing strategy that equally invests in all the significant anomalies as described in Figure 2. The Panel B shows transaction costs for the shrinkage strategy described in Table VI that is estimated on individual stock's returns from the US. The transaction costs are estimated with VoV(% Spread) proxy of Fong et al. (2017).

asset managers. We will therefore focus only on the fixed component (effective spread). Fong et al. (2017) benchmark the proxy to other existing proxies and find that it can only be outperformed by quoted spread, which is, however, not available for all the regions over our whole sample period.

It is evident that the lowest trading costs are indeed in the US. The highest are then in emerging part Asia Pacific. The peaks in figure appear every July because of annual rebalancing of the fundamental strategies. The graph also documents that there are periods with significant spillover of illiquidity. Two such major episodes are Financial Crisis of 2008 and dot com bubble of early 2000s.

Table XI presents returns on the mixing strategy introduced in III adjusted for the trading costs. Returns in all of the region, with the exception of Asia Pacific Emerging, are insignificant

Table XI
Performance of portfolio mixing strategies after transaction costs

The table shows returns minus transaction costs of the strategy described in Table III. The transaction costs are estimated with VoV(% Spread) proxy of Fong et al. (2017).

	July 1990+					July 2000+	
	USA	E	J	AP	Global	CN	AP EM
Panel A: Equal-weighted portfolios							
Mean Return	0.22	0.19	-0.02	0.12	0.13	0.10	0.30
	(1.53)	(1.36)	(-0.17)	(0.90)	(1.31)	(1.16)	(2.49)
CAPM alpha	0.34	0.22	-0.02	0.15	0.18	0.10	0.33
	(2.21)	(1.56)	(-0.17)	(1.12)	(1.82)	(1.15)	(2.96)
FF3 alpha	0.17	0.08	-0.12	-0.00	0.04	0.05	0.31
	(2.42)	(1.13)	(-1.60)	(-0.02)	(0.75)	(0.67)	(2.63)
FF5 alpha	-0.01	-0.12	-0.12	-0.14	-0.04	0.10	0.29
	(-0.10)	(-1.41)	(-1.70)	(-1.12)	(-0.85)	(1.17)	(2.65)
Panel B: Value-weighted portfolios							
Mean Return	-0.04	0.07	0.10	-0.08	0.01	0.03	0.06
	(-0.24)	(0.64)	(0.85)	(-0.28)	(0.12)	(0.28)	(0.44)
CAPM alpha	0.09	0.11	0.11	-0.02	0.05	0.03	0.09
	(0.54)	(0.90)	(0.84)	(-0.06)	(0.45)	(0.27)	(0.60)
FF3 alpha	0.01	0.01	0.15	-0.02	-0.01	-0.01	0.10
	(0.05)	(0.05)	(1.06)	(-0.07)	(-0.08)	(-0.13)	(0.70)
FF5 alpha	-0.19	-0.14	0.15	-0.28	-0.09	0.07	0.10
	(-1.21)	(-0.87)	(1.09)	(-0.98)	(-0.75)	(0.69)	(0.80)

after the transaction costs. This gives a simple explanation for higher profitability of the individual anomalies outside the US in that arbitrage is more costly there. It thus does not pay to trade on these signals individually although they seemingly appear profitable.

C.2. Shrinkage strategies

We proceed with the same methodology applied to our shrinkage strategies. Panel B in Figure 5 maps transaction costs on our gradient boosting strategy estimated in the US. It is apparent that transaction costs are much higher than in the case of individual anomalies. This is because a large portion of the individual anomalies were fundamental anomalies that are rebalanced annually whereas the shrinkage strategies are rebalanced monthly. The transaction costs have decreased significantly over time and there are again several historical episodes where they were heavily elevated, one being the financial crisis of 2008. They are much lower on value-weighted portfolios relative to equal-weighting which is expected as value-weighting puts larger weight on more liquid stocks. The costs are again the smallest in the US.

Table XII presents mean returns on the shrinkage strategies after transaction costs. They are mostly insignificant at 5% level for least squares shrinkage regressions but the returns remain positive. The mean returns on gradient boosting shrinkage are, however, significantly positive. The net return in the US are close to 20% a year. The returns elsewhere are lower due to the higher transaction costs there. Panel B then presents net mean returns after transaction costs for universe of stocks with trading volume in the previous month larger than $100 million. These stocks should be very liquid and with virtually no fixed transaction costs. It should also be possible to invest large quantities of money in them without huge price impact. The net mean returns for least square shrinkage regressions are again insignificant but returns on gradient boosting shrinkage remain significantly positive. This documents that choice of appropriate forecasting method is very important for success of investing into the anomalies.

Table XII
Performance of shrinkage strategies after transaction costs

The table shows mean returns after transaction costs of the shrinkage strategy described in Table VI that is estimated on the individual stocks from the US. The transaction costs are estimated with VoV(% Spread) proxy of Fong et al. (2017). We further restrict our large cap universe of stocks to those with dollar trading volume of at least $100 million in Panel B. We adjust standard error for heteroskedasticity and autocorrelation up to 12 lags with Newey and West (1987) adjustment. The returns are in percentage points per month.

	Equal-weighted				Value-weighted			
	USA	Europe	Japan	Asia Pacific	USA	Europe	Japan	Asia Pacific
Panel A: Full large cap universe								
WLS								
Mean Return	0.925	0.556	0.205	0.488	0.664	0.678	0.499	0.647
	(1.850)	(1.430)	(0.775)	(1.470)	(1.440)	(2.070)	(1.580)	(1.810)
FF5 alpha	0.374	-0.158	-0.023	0.207	0.179	0.062	0.294	0.226
	(1.430)	(-0.671)	(-0.094)	(0.586)	(0.558)	(0.258)	(0.934)	(0.590)
Gradient boosting regression trees								
Mean Return	1.400	0.819	0.507	0.836	1.600	0.756	0.371	1.230
	(3.090)	(2.340)	(2.270)	(2.250)	(3.820)	(2.420)	(1.300)	(2.880)
FF5 alpha	0.931	0.331	0.317	0.655	1.340	0.317	0.180	0.681
	(3.510)	(1.470)	(1.530)	(1.740)	(4.330)	(1.220)	(0.677)	(1.970)
Panel B: Stocks with volume in the previous month larger than $100 million								
WLS								
Mean Return	0.630	0.648	0.072	0.222	0.535	0.516	0.216	0.116
	(1.350)	(1.780)	(0.159)	(0.582)	(1.160)	(1.700)	(0.554)	(0.318)
FF5 alpha	0.167	0.009	-0.159	-0.546	0.012	0.020	0.003	-0.705
	(0.695)	(0.036)	(-0.352)	(-0.994)	(0.037)	(0.077)	(0.008)	(-1.390)
Gradient boosting regression trees								
Mean Return	1.020	1.160	0.803	0.924	1.400	0.848	0.300	1.010
	(3.150)	(3.500)	(2.270)	(1.910)	(4.140)	(2.960)	(0.722)	(2.010)
FF5 alpha	0.716	0.629	0.663	0.743	1.400	0.467	0.159	0.458
	(2.430)	(2.640)	(1.700)	(1.360)	(4.180)	(1.550)	(0.323)	(0.885)
Diff wrt OLS	0.388	0.511	0.731	0.702	0.860	0.332	0.084	0.895
	(1.340)	(2.660)	(1.900)	(1.400)	(1.880)	(1.250)	(0.200)	(2.210)

D. Post publication decay due to informed trading

D.1. Post publication decay on individual anomalies in the US

McLean and Pontiff (2016) have shown that returns on individual anomalies drop out-of-sample of original studies and post publication. We start our analysis with replication of their key findings on our universe of anomalies and for our differently constructed portfolios. McLean and Pontiff (2016) have introduced an ingenious identification scheme to measure the size of data mining and profitability lost due to post publication trading. The estimation of size of data mining is based on a simple idea that during the period between publication of a study and the end of sample period in it, any drop in performance should be due to data mining. The performance tends to decrease further after publication and this decay is due to market participants trading on the anomalies. Harvey et al. (2016) showed that results of many academic studies are due to data mining when considered in multiple hypothesis testing framework. It is thus important to understand out-of-sample and post publication decay to measure the amount of data mining in finance academia to propose appropriate adjustments.[11] Knowing the expected decay in performance can then help us to understand the previous results on value of past performance for future predictions. This is important for anyone trying to apply the anomalies research in practice.

Table XIII presents results for the following regression:

$$r_{it} = \alpha_i + \beta_1 Post\ Sample\ dummy_{it} + \beta_2 Post\ Publication\ dummy_{it} + \epsilon_{it} \tag{5}$$

for various restrictions on sample of anomalies and different constructions of underlying portfolios. Portfolios in original weighting and value-weighted groups are constructed on full sample of stocks. The third group then excludes all stocks with capitalization lower than the bottom decile of NYSE at the end of previous June. There are 3 groups of anomalies: "all" category includes all our 153 anomalies, "signif" category follows McLean and Pontiff (2016) and includes those significant on full sample with original weighting of returns, and finally "signif 2" category those that are significant on large cap universe with value-weighted returns.[12] The standard errors in regressions are robust to general forms of spatial and temporal dependence as in Driscoll and Kraay (1998).

The results with equal weighting are similar to those in McLean and Pontiff (2016) in that the full decay after publication cannot be explained by data mining. The size of decay is slightly different but this can be due to sampling variation. Value-weighted returns on full universe of stocks already paint a different picture in that the full out-of-sample decay appears to be due to data mining for a set of strategies that are the most profitable on investable universe. Value-weighted returns on the large cap universe are very similar to this. The larger decay for large caps and value-weighted returns can be explained by lower transaction costs and generally lower arbitrage costs there. Anyone who would want to profit from the anomalies would start with investments there. The fact that full decay already happens before publication can also be explained by some early adopter who can observe the results before publication. This is very likely true in many cases as some studies were published by researchers from quantitative asset managers.[13] But if this was the case then these asset managers tend to have global presence and the same decay should appear in other regions as well. The next section will focus precisely on this.

D.2. Post publication decay on individual anomalies around the globe

The previous section has confirmed that there is a strong post-publication decay in our sample of anomalies. We will now study how much of the decay is due to data mining by adding international evidence. Similar analysis has already been done by Jacobs and Müller (2017b) we thus here replicate their main finding for our sample of anomalies. Our analysis is also an extension in that we focus on investable large cap universe of stocks with value-weighted returns and it is thus closer to what investors should expect if they want to trade on the anomalies. If all of the decay in the US is due to data mining then there should be no decay around the publication of the anomalies in any other region. International evidence offers a viable out-of-sample test as the authors of the studies were not looking for the new anomalies there. The international evidence was published later in many cases but this should have no effect on the anomalies identified in the original studies. The international evidence can also help to uncover the component of the decay that is due to trading of quantitative asset managers. The quantitative investors have to be very sophisticated and must build up an expensive infrastructure to reduce price impact of their trades. It is thus very unlikely

[11] Out-of-sample indicates a period that begins after the end of sample in the original study.

[12] We select significant anomalies on the same sample as in the original studies including even the period before 1963.

[13] See, for example, Frazzini and Pedersen (2014) and many other studies coming from researchers affiliated with AQR.

Table XIII
Post publication decay in the US

The table shows panel regressions of returns on anomalies regressed on dummy variables for the period between the end of sample and publication (Post Sample) and period post publication of the anomalies. We construct portfolios for the anomalies in three ways: using the weighting of returns and sample as in the original paper, using value-weighted returns and sample as in the original paper, and using value-weighted returns and sample that excludes stocks with capitalization smaller than bottom decile of NYSE at the end of previous June. The sample spans July 1963 to December 2016. The differently weighted portfolios are further grouped into all available anomalies, anomalies that are significant using the original weighting, and anomalies that are significant with value-weighting at large cap universe. We use anomalies fixed effects for all regressions and adjust the standard errors for heteroskedasticity and autocorrelation as in Driscoll and Kraay (1998).

	Original Weighting			Value-weighted			Value-weighted Large Cap		
	All	Signif	Signif 2	All	Signif	Signif 2	All	Signif	Signif 2
Intercept	0.59	0.82	0.89	0.44	0.61	0.79	0.38	0.51	0.67
	(11.70)	(12.60)	(13.30)	(10.50)	(9.82)	(11.40)	(9.64)	(8.27)	(9.83)
Post Sample	-0.08	-0.08	-0.30	-0.17	-0.20	-0.48	-0.17	-0.20	-0.44
	(-0.88)	(-0.53)	(-2.19)	(-1.86)	(-1.44)	(-3.60)	(-2.04)	(-1.51)	(-3.48)
Post Publication	-0.27	-0.40	-0.50	-0.27	-0.37	-0.51	-0.20	-0.30	-0.44
	(-3.25)	(-3.85)	(-4.77)	(-3.26)	(-3.28)	(-3.99)	(-2.58)	(-2.73)	(-3.48)
Anomalies Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Sample Size	93665	57616	34069	93665	57616	34069	93665	57616	34069

that they would leave out a large fraction of global market and focus solely on the US. Their activity in the other regions would then depress the returns there which would show up in our analysis.

Table XIV provides regressions that are similar in nature to those in Table XIII. All the regressions are some variation of:

$$
\begin{aligned}
r_{it} = \alpha_i + \beta_1 Post\ Sample\ dummy_{it} + \beta_2 Post\ Publication\ dummy_{it} \\
+\beta_3 Post\ Sample\ dummy\ USA_{it} + \beta_4 Post\ Publication\ dummy\ USA_{it} + \epsilon_{it}\ .
\end{aligned}
$$

We study countries in Asia Pacific separately for sake of robustness as there is only small fundamental coverage there before 2000. The overall number of liquid companies is also much lower there with respect to the other regions. The table is based on portfolios that start in July 1990. We include anomalies fixed effects in most of the specifications. We include only anomalies that are significant with t-statistic larger than 1.96 at the time of their publication. We exclude all anomalies with in-sample period ending before July 1992 so that the in-sample period is at least two years. This gives together about 16 000 observations in each of the regions. We adjust the standard errors for heteroskedasticity and autocorrelation as in Driscoll and Kraay (1998).

The table provides very similar results in the US as for the longer sample in the Table XIII. The decay in returns happens immediately out-of-sample and even before the publication. Other regions paint very different picture. The decay is not significant in all of them although there is some decay in Europe. This could be explained by high correlation of strategies between Europe and the US or, alternatively, the quantitative asset managers are investing there more than in Japan. Pooling all the regions together offers very similar picture. The post sample and post publication dummies in the US siphon out all the decay. The results are robust to inclusion of anomalies fixed effects. This goes sharply against the hypothesis that the post publication decay is due to informed trading. The anomalies are at least as much profitable outside the US as in the

Table XIV
Post publication decay around the globe

The table shows panel regressions of returns on anomalies regressed on dummy variables for the period between end of sample and publication (Post Sample) and period post publication of the anomalies in four global regions. The sample spans July 1990 to December 2016 and excludes stocks with capitalization smaller than bottom decile of NYSE at the end of previous June. The table is based on all anomalies that are significant in-sample of the original studies at 5% level with the same weighting and sample restrictions as for the portfolios. We use anomalies fixed effects for regressions in individual regions and adjust the standard errors for heteroskedasticity and autocorrelation as in Driscoll and Kraay (1998).

	USA	E	J	AP	Excluding AP		All		
	I.	II.	III.	IV.	V.	VI.	VII.	VIII.	IX.
Intercept	0.75	0.52	0.04	0.52	0.60	0.75	0.59	0.75	0.47
	(6.71)	(4.23)	(0.28)	(2.16)	(5.70)	(6.93)	(5.34)	(6.93)	(3.90)
Post Sample	-0.56	-0.16	0.13	0.03	-0.24	-0.06	-0.21	-0.08	-0.00
	(-4.05)	(-1.06)	(0.60)	(0.11)	(-2.14)	(-0.49)	(-1.72)	(-0.56)	(-0.01)
Post Publication	-0.53	-0.14	0.18	0.02	-0.21	-0.04	-0.18	-0.05	0.02
	(-3.38)	(-0.93)	(0.87)	(0.07)	(-1.69)	(-0.25)	(-1.29)	(-0.32)	(0.10)
Post Sample USA						-0.50		-0.48	-0.56
						(-3.55)		(-3.46)	(-3.76)
Post Publication USA						-0.49		-0.47	-0.55
						(-3.50)		(-3.13)	(-3.16)
J					-0.35	-0.59	-0.36	-0.59	
					(-3.17)	(-4.37)	(-3.18)	(-4.21)	
E					-0.05	-0.29	-0.05	-0.28	
					(-0.65)	(-2.70)	(-0.68)	(-2.62)	
AP							0.05	-0.18	
							(0.34)	(-1.06)	
Anomalies Fixed Effects	YES	YES	YES	YES					YES
Sample Size	18408	16037	16031	15771	50476	50476	66247	66247	66247

US and this should attract attention of arbitrageurs. To summarize, all of the out-of-sample decay documented so far can be attributed to data mining in the US. We have found only very small support for the hypothesis that it is due to trading of quantitative asset managers.

D.3. Can decrease in profitability of individual anomalies explain the decay?

The previous regressions have not accounted for changing overall profitability of all the strategies over time. Green et al. (2017) showed that returns on anomalies have decreased after 2003. Why should the markets be more efficient in the recent period? First, financial data is much more readily available to a larger pool of investors. There are now many websites that specialize on financial news and offer financial data for free. High quality fundamental data used to be a privilege of select few. Computing power has also increased exponentially every year following Moore's law. Finally, the transaction costs have come down significantly. This calls to question whether the out-of-sample decay in returns is not in fact capturing this overall increase in market efficiency.

Table XV adds a dummy for time period after 2003 to regressions in Tables XIII and XIV. The specifications I. and II. use longer sample beginning in July 1963. All the regressions are based on value-weighted returns from large cap universe with the exception of the specification I. which is based on full sample of stocks with returns weighted as in the original studies. This is consistent with McLean and Pontiff (2016). The addition of dummy variable for 2003+ periods destroys the out-of-sample decay presented so far in all of the specifications.[14] This leads us to conclude

[14]This is in contrast with McLean and Pontiff (2016) who show that their results are robust to time effects. Some of this difference can be explained by our more recent sample and larger set of anomalies. We also use a different adjustment of standard errors that is more conservative than theirs as it also controls for autocorrelation. Our

Table XV
Role of time effects for post publication decay

The table adds a dummy variable for time period after 2003 in regressions presented in Tables XIII and XIV. See Table XIV for description of the regressions. Specification I. uses the original weighting of the portfolios with sample start in July 1963. Other specifications use value-weighted returns on large cap universe with sample start in July 1990 with the exception of II. which starts in July 1963.

	USA			E	J	AP	All		
	I.	II.	III.	IV.	V.	VI.	VII.	VIII.	IX.
Intercept	0.83	0.68	0.84	0.61	0.03	0.59	0.65	0.85	0.54
	(12.70)	(9.94)	(6.49)	(4.29)	(0.16)	(2.10)	(4.99)	(6.59)	(4.02)
Post Sample	0.20	-0.01	0.09	0.52	0.11	0.94	0.09	0.25	0.59
	(1.11)	(-0.03)	(0.35)	(1.88)	(0.38)	(3.00)	(0.54)	(1.48)	(2.73)
Post Publication	0.06	0.24	0.46	0.82	0.26	1.33	0.29	0.47	0.89
	(0.31)	(0.96)	(1.64)	(2.90)	(0.84)	(3.49)	(1.83)	(2.64)	(3.71)
2003+	-0.52	-0.76	-1.07	-1.00	-0.04	-1.25	-0.36	-0.55	-0.64
	(-2.68)	(-2.94)	(-3.16)	(-2.99)	(-0.14)	(-2.82)	(-2.98)	(-3.90)	(-4.06)
Post Sample USA								-0.56	-0.64
								(-3.26)	(-3.17)
Post Publication USA								-0.65	-0.85
								(-4.07)	(-3.12)
J							-0.02	-0.31	
							(-0.28)	(-2.49)	
E							0.08	-0.20	
							(0.54)	(-1.01)	
AP							-0.54	-0.55	
							(-2.27)	(-2.30)	
Anomalies Fixed Effects	YES	YES	YES	YES	YES	YES			YES
Sample Size	57616	34069	16992	13736	13730	13470	57928	57928	57928

that the out-of-sample decay has indeed much to do with overall shift in profitability of all the anomalies. It nevertheless does not affect our previous conclusion that the individual anomalies are less profitable post publication.

D.4. Is there any decay on the shrinkage strategy?

The drop in profitability of all anomalies in the US is, however, related only to disaggregated analysis of the individual anomalies. Returns on gradient boosting shrinkage strategy in the US in Figure 4 have not suffered from such a drop and are comparable to the other regions. The whole analysis could thus depend on the specific methodology that emphasizes linear relationships. We further test this hypothesis by adapting our regressions to the shrinkage GBRT strategy in Table XVI. We substitute portfolios on individual anomalies by portfolios on the shrinkage strategy that is estimate on all the anomalies published before either 1995, 2000, 2005, or 2010. The regressions should thus measure non-linear explanatory power of all anomalies after their publication.

The returns on the strategy again decay only in the US. This is due to larger in-sample returns there and this decay now merely equalizes out-of-sample returns in all the regions. The higher in-sample returns are likely due to in-sample overfitting as all the anomalies were discovered in the US. All the results again do not survive inclusion of time effects in the regressions and the decay can then easily be due to changes in profitability of all the strategies over time.

results are, nonetheless, consistent with Jacobs and Müller (2017b).

Table XVI
Is there any decay on the shrinkage strategy?

The table shows panel regressions of value-weighted returns of GBRT shrinkage strategies described in Table VI that are estimated on the individual stocks from the US. We restrict anomalies in the estimation to those that were published before 1995, 2000, 2005, or 2010 and we thus have a panel with returns on four portfolios. We then regress the returns on dummy variable for the period post publication of the anomalies (e.g. 1995+, 2000+...). The sample spans July 1990 to December 2016 and excludes stocks with capitalization smaller than bottom decile of NYSE at the end of previous June. We use strategy fixed effects for regressions in the individual regions and adjust the standard errors for heteroskedasticity and autocorrelation as in Driscoll and Kraay (1998).

	USA	E	J	AP	All		USA	All
	I.	II.	III.	IV.	V.	VI.	VII.	VIII.
Intercept	2.42	1.52	1.95	1.52	2.08	2.53	2.77	2.96
	(7.16)	(5.15)	(5.73)	(3.88)	(6.71)	(7.57)	(7.32)	(7.10)
Post Publication	-1.12	-0.10	-0.58	0.62	-0.48	-0.20	0.96	0.39
	(-3.24)	(-0.32)	(-1.70)	(1.45)	(-2.29)	(-0.88)	(1.82)	(1.54)
Post Publication USA						-1.12		-0.36
						(-3.76)		(-0.85)
J					-0.18	-0.77		-1.02
					(-0.52)	(-2.12)		(-2.34)
E					-0.36	-0.96		-1.20
					(-1.26)	(-2.74)		(-2.93)
AP					0.02	-0.58		-0.82
					(0.06)	(-1.82)		(-2.13)
2003+							-2.75	-0.95
							(-3.53)	(-2.02)
2003+ USA								-1.22
								(-1.95)
Strategy Fixed Effects	YES	YES	YES	YES			YES	
Sample Size	1272	1272	1272	1272	5088	5088	1272	5088

VI. Conclusion

We have studies profitability of quantitative strategies based on published anomalies around the globe. We have shown that investing into individual anomalies is not profitable after accounting for transaction costs and it is important to combine multiple signals with appropriate methods. Machine learning methods lead to higher (risk adjusted) returns relative to standard methods applied in academic finance literature. The strategies are then profitable even on large cap universe of stocks that are highly liquid. We also document that new anomalies improve average returns on our investment strategy after accounting for the previously published anomalies. New anomalies studies are thus successful in finding new sources of risk and behavioural biases.

It is customary in the academic literature to assume that anomalies that have historically worked everywhere are truly driven by important risk factors and this guarantees that they will persist into the future. We show that this does not have to be the case and past performance in international markets is not a good predictor of future returns of the individual quantitative strategies. We come to the same conclusion with more sophisticated machine learning methods where out-of-sample performance in the US is not improved with inclusion of international evidence. Predictions of future returns fitted in the US possess most of information about the expected stock returns in all the other regions.

Appendix A. Adjustments in Datastream

We apply a series of adjustments on the raw returns to improve their quality. We require that return index (RI) is larger than 0.001 on the first day of the month for precision reason. we set RI to missing if daily return is larger than 500% or if price on the first day of the month is larger than $1 million. We also set as missing any monthly return larger than 2000%. Datastream provides stale prices when there is no trade during the day or when the stock is no longer traded so that the price of the last trade is repeated until new information arrives. We thus delete the latest observations of price with no trading. Following Tobek and Hronec (2018) we fix daily returns when there are stale price quotes around corporate events. Following Ince and Porter (2006) we also set as missing monthly returns over 300% that revert back over the next month.[15] We winsorize .01% of returns in each region and year before 2000.

Appendix B. List of anomalies

Table XVII
List of Anomalies

Fundamental	
Accruals	
Accruals	Sloan (1996)
Change in Common Equity	Richardson, Sloan, Soliman, and Tuna (2006)
Change in Current Operating Assets	Richardson et al. (2006)
Change in Current Operating Liabilities	Richardson et al. (2006)
Change in Financial Liabilities	Richardson et al. (2006)
Change in Long-Term Investments	Richardson et al. (2006)
Change in Net Financial Assets	Richardson et al. (2006)
Change in Net Non-Cash Working Capital	Richardson et al. (2006)
Change in Net Non-Current Operating Assets	Richardson et al. (2006)
Change in Non-Current Operating Assets	Richardson et al. (2006)
Change in Non-Current Operating Liabilities	Richardson et al. (2006)
Change in Short-Term Investments	Richardson et al. (2006)
Discretionary Accruals	Dechow, Sloan, and Sweeney (1995)
Growth in Inventory	Thomas and Zhang (2002)
Inventory Change	Thomas and Zhang (2002)
Inventory Growth	Belo and Lin (2011)
M/B and Accruals	Bartov and Kim (2004)
Net Working Capital Changes	Soliman (2008)
Percent Operating Accrual	Hafzalla, Lundholm, and Matthew Van Winkle (2011)
Percent Total Accrual	Hafzalla et al. (2011)
Total Accruals	Richardson et al. (2006)
Intangibles	
$\triangle$ Gross Margin - $\triangle$ Sales	Abarbanell and Bushee (1998)
$\triangle$ Sales - $\triangle$ Accounts Receivable	Abarbanell and Bushee (1998)
$\triangle$ Sales - $\triangle$ Inventory	Abarbanell and Bushee (1998)
$\triangle$ Sales - $\triangle$ SG and A	Abarbanell and Bushee (1998)
Asset Liquidity	Ortiz-Molina and Phillips (2014)
Asset Liquidity II	Ortiz-Molina and Phillips (2014)
Cash-to-assets	Palazzo (2012)
Earnings Conservatism	Francis, LaFond, Olsson, and Schipper (2004)
Earnings Persistence	Francis et al. (2004)
Earnings Predictability	Francis et al. (2004)

[15] Specifically, we set as missing returns in two consecutive months if the return in the first was larger than 300% and the overall return over the two months was lower than 50%.

Earnings Smoothness	Francis et al. (2004)
Earnings Timeliness	Francis et al. (2004)
Herfindahl Index	Hou and Robinson (2006)
Hiring rate	Belo, Lin, and Bazdresch (2014)
Industry Concentration Assets	Hou and Robinson (2006)
Industry Concentration Book Equity	Hou and Robinson (2006)
Industry-adjusted Organizational Capital-to-Assets	Eisfeldt and Papanikolaou (2013)
Industry-adjusted Real Estate Ratio	Tuzel (2010)
Org. Capital	Eisfeldt and Papanikolaou (2013)
RD / Market Equity	Chan, Lakonishok, and Sougiannis (2001)
RD Capital-to-assets	Li (2011)
RD Expenses-to-sales	Chan et al. (2001)
Tangibility	Hahn and Lee (2009)
Unexpected RD Increases	Eberhart, Maxwell, and Siddique (2004)
Whited-Wu Index	Whited and Wu (2006)
Investment	
$\triangle$ CAPEX - $\triangle$ Industry CAPEX	Abarbanell and Bushee (1998)
Asset Growth	Cooper, Gulen, and Schill (2008)
Change Net Operating Assets	Hirshleifer, Hou, Teoh, and Zhang (2004)
Changes in PPE and Inventory-to-Assets	Lyandres, Sun, and Zhang (2007)
Composite Debt Issuance	Lyandres et al. (2007)
Composite Equity Issuance (5-Year)	Daniel and Titman (2006)
Debt Issuance	Spiess and Affleck-Graves (1995)
Growth in LTNOA	Fairfield, Whisenant, and Yohn (2003)
Investment	Titman, Wei, and Xie (2004)
Net Debt Finance	Bradshaw, Richardson, and Sloan (2006)
Net Equity Finance	Bradshaw et al. (2006)
Net Operating Assets	Hirshleifer et al. (2004)
Noncurrent Operating Assets Changes	Soliman (2008)
Share Repurchases	Ikenberry, Lakonishok, and Vermaelen (1995)
Total XFIN	Bradshaw et al. (2006)
Profitability	
Asset Turnover	Soliman (2008)
Capital Turnover	Haugen and Baker (1996)
Cash-based Operating Profitability	Ball, Gerakos, Linnainmaa, and Nikolaev (2016)
Change in Asset Turnover	Soliman (2008)
Change in Profit Margin	Soliman (2008)
Earnings / Price	Basu (1977)
Earnings Consistency	Alwathainani (2009)
F-Score	Piotroski (2000)
Gross Profitability	Novy-Marx (2013)
Labor Force Efficiency	Abarbanell and Bushee (1998)
Leverage	Bhandari (1988)
O-Score (More Financial Distress)	Dichev (1998)
Operating Profits to Assets	Ball et al. (2016)
Operating Profits to Equity	Fama and French (2015)
Profit Margin	Soliman (2008)
Return on Net Operating Assets	Soliman (2008)
Return-on-Equity	Haugen and Baker (1996)
Z-Score (Less Financial Distress)	Dichev (1998)
Value	
Assets-to-Market	Fama and French (1992)
Book Equity / Market Equity	Fama and French (1992)
Cash Flow / Market Equity	Lakonishok, Shleifer, and Vishny (1994)
Duration of Equity	Dechow, Sloan, and Soliman (2004)
Enterprise Component of Book/Price	Penman, Richardson, and Tuna (2007)
Enterprise Multiple	Loughran and Wellman (2011)
Intangible Return	Daniel and Titman (2006)
Leverage Component of Book/Price	Penman et al. (2007)

Net Payout Yield	Boudoukh, Michaely, Richardson, and Roberts (2007)
Operating Leverage	Novy-Marx (2010)
Payout Yield	Boudoukh et al. (2007)
Sales Growth	Lakonishok et al. (1994)
Sales/Price	Barbee Jr, Mukherji, and Raines (1996)
Sustainable Growth	Lockwood and Prombutr (2010)
Frictions	
11-Month Residual Momentum	Blitz, Huij, and Martens (2011)
52-Week High	George and Hwang (2004)
Amihud's Measure (Illiquidity)	Amihud (2002)
Beta	Fama and MacBeth (1973)
Betting against Beta	Frazzini and Pedersen (2014)
Bid-Ask Spread	Amihud and Mendelson (1986)
Cash Flow Variance	Haugen and Baker (1996)
Coefficient of Variation of Share Turnover	Chordia, Subrahmanyam, and Anshuman (2001)
Coskewness	Harvey and Siddique (2000)
Downside Beta	Ang, Chen, and Xing (2006a)
Earnings Forecast-to-Price	Elgers, Lo, and Pfeiffer Jr (2001)
Firm Age	Barry and Brown (1984)
Firm Age-Momentum	Zhang (2006)
Idiosyncratic Risk	Ang, Hodrick, Xing, and Zhang (2006b)
Industry Momentum	Moskowitz and Grinblatt (1999)
Lagged Momentum	Novy-Marx (2012)
Liquidity Beta 1	Acharya and Pedersen (2005)
Liquidity Beta 2	Acharya and Pedersen (2005)
Liquidity Beta 3	Acharya and Pedersen (2005)
Liquidity Beta 4	Acharya and Pedersen (2005)
Liquidity Beta 5	Acharya and Pedersen (2005)
Liquidity Shocks	Bali, Peng, Shen, and Tang (2013)
Long-Term Reversal	Bondt and Thaler (1985)
Max	Bali et al. (2011)
Momentum	Jegadeesh and Titman (1993)
Momentum and LT Reversal	Kot and Chan (2006)
Momentum-Reversal	Jegadeesh and Titman (1993)
Momentum-Volume	Lee and Swaminathan (2000)
Price	Blume and Husic (1973)
Seasonality	Heston and Sadka (2008)
Seasonality 1 A	Heston and Sadka (2008)
Seasonality 1 N	Heston and Sadka (2008)
Seasonality 11-15 A	Heston and Sadka (2008)
Seasonality 11-15 N	Heston and Sadka (2008)
Seasonality 16-20 A	Heston and Sadka (2008)
Seasonality 16-20 N	Heston and Sadka (2008)
Seasonality 2-5 A	Heston and Sadka (2008)
Seasonality 2-5 N	Heston and Sadka (2008)
Seasonality 6-10 A	Heston and Sadka (2008)
Seasonality 6-10 N	Heston and Sadka (2008)
Share Issuance (1-Year)	Pontiff and Woodgate (2008)
Share Turnover	Datar, Naik, and Radcliffe (1998)
Short-Term Reversal	Jegadeesh (1990)
Size	Banz (1981)
Tail Risk	Kelly and Jiang (2014)
Total Volatility	Ang et al. (2006b)
Volume / Market Value of Equity	Haugen and Baker (1996)
Volume Trend	Haugen and Baker (1996)
Volume Variance	Chordia et al. (2001)
I/B/E/S	

Analyst Value	Frankel and Lee (1998)
Analysts Coverage	Elgers et al. (2001)
Change in Forecast + Accrual	Barth and Hutton (2004)
Change in Recommendation	Jegadeesh, Kim, Krische, and Lee (2004)
Changes in Analyst Earnings Forecasts	Hawkins, Chamberlin, and Daniel (1984)
Disparity between LT and ST Earnings Growth Forecasts	Da and Warachka (2011)
Dispersion in Analyst LT Growth Forecasts	Anderson, Ghysels, and Juergens (2005)
Down Forecast	Barber, Lehavy, McNichols, and Trueman (2001)
Forecast Dispersion	Diether, Malloy, and Scherbina (2002)
Long-Term Growth Forecasts	La Porta (1996)
Up Forecast	Barber et al. (2001)

Online Appendix

Anomalies are grouped into 5 categories: accruals, profitability, value, investment, an intangibles. Construction of individual anomalies follows Harvey et al. (2016), McLean and Pontiff (2016) and Hou et al. (2017), with the exception of selecting subset of exchanges and frequency of rebalancing. When these exceptions apply, they are described in individual anomalies definitions.

FUNDAMENTAL

Accruals

Accruals (Acc)

Based on Sloan (1996), accruals are defined as

$$Acc = \frac{(\Delta act_t - \Delta che_t) - (\Delta lct_t - \Delta dlc_t - \Delta tp_t) - dp_t}{(at_t + at_{t-1})/2}$$

where Δact_t is change in current assets, Δche_t is change in cash and cash equivalents, Δlct_t is annual change in current liabilities, Δdlc_t is annual change in debt included in current liabilities, Δtp_t is annual change in income taxes payable and dp is depreciation and amortization expense.

Change in Current Operating Assets (ChCOA)

Based on Richardson et al. (2006), change in current operating assets is defined as

$$ChCOA = \frac{COA_t - COA_{t-1}}{at_{t-1}}$$

where COA_t are current operating assets, $COA_t = act_t - che_t$ in which act_t are current assets, che_t are cash and short-term investment and at_{t-1} are one-year lagged total assets

Change in Current Operating Liabilities (ChCOL)

Based on Richardson et al. (2006), change in current operating liabilities is defined as

$$ChCOL = \frac{COL_t - COL_{t-1}}{at_{t-1}}$$

where COL_t are current operating liabilities, $COL_t = lct_t - dlc_t$ in which lct_t are current liabilities, dlc_t is debt in current liabilities and at_{t-1} are one-year lagged total assets.

Change in Net Non-Cash Working Capital (ChNNCWC)

Based on Richardson et al. (2006), Change in Net Non-Cash Working Capital is defined as

$$ChNNCWC = \frac{WC_t - WC_{t-1}}{at_{t-1}}$$

where WC_t is working capital, $WC_t = COA_t - COL_t$ in which COA_t are current operating assets defined above in Change in Current Operating Assets anomaly and COL_t are current operating liabilities defined above in Change in Current Operating Liabilities anomaly.

Change in Net Non-Current Operating Assets (ChNNCOA)

Based on Richardson et al. (2006), Change in Net Non-Current Operating Assets is defined as

$$ChNNCOA = \frac{NCOA_t - NCOA_{t-1}}{at_{t-1}}$$

where NCO_t are non-current operating asset, $NCOA_t = NCA_t - NCL_t$ in which NCA_t are non-current assets defined in Change in Non-Current Operating Assets anomaly and NCL_t are non-current operating liabilities defined in Change in Non-Current Operating Liabilities anomaly.

Change in Non-Current Operating Assets (ChNCOA)

Based on Richardson et al. (2006), Change in Non-Current Operating Assets is defined as

$$ChNCOA = \frac{NCA_t - NCA_{t-1}}{at_{t-1}}$$

where NCA_t are non-current assets defined as $NCA_t = at_t - act_t - ivao_t$ where at_t are total assets, act_t are current assets, $ivao_t$ is investment and advances (0 if missing).

Change in Non-Current Operating Liabilities (ChNCOL)

Based on Richardson et al. (2006), Change in Non-Current Operating Liabilities is defined as

$$ChNCOL = \frac{NCL_t - NCL_{t-1}}{at_{t-1}}$$

where $NCL_t = lt_t - lct_t - dltt_t$ in which lt_t are total liabilities, lct_t are current liabilities and $dltt_t$ is long-term debt (0 if missing).

Change in Net Financial Assets (ChNFA)

Based on Richardson et al. (2006), Change in Net Financial Assets is defined as

$$ChNFA = \frac{NFNA_t - NFNA_{t-1}}{at_{t-1}}$$

where

$$NFNA_t = FNA_t - FNL_t$$

are net financial assets in which FNA_t are financial assets, $FNA_t = ivst_t + ivao_t$ where $ivst_t$ are short-term investments, $ivao_t$ are long-term investments and FNL_t are financial liabilities, $FNL_t = dltt_t + dlc_t + pstk_t$ in which $dltt_t$ is long-term debt, dlc_t is debt in current liabilities and $pstk_t$ is preferred stock.

Change in Short-Term Investments (ChSTI)

Based on Richardson et al. (2006), Change in Short-Term Investments is defined as

$$ChSTI = \frac{ivst_t - ivst_{t-1}}{at_{t-1}}$$

where $ivst_t$ are short-term investments and at_{t-1} are one-year lagged total assets.

Change in Long-Term Investments (ChLTI)

Based on Richardson et al. (2006), Change in Long-Term Investments is defined as

$$ChLTI = \frac{ivao_t - ivao_{t-1}}{at_{t-1}}$$

where $ivao_t$ are long-term investments and at_{t-1} are one-year lagged total assets.

Change in Common Equity (ChCE)

Based on Richardson et al. (2006), Change in Common Equity is defined as

$$ChCE = \frac{ceq_t - ceq_{t-1}}{at_{t-1}}$$

where ceq_t is common equity and at_{t-1} are one-year lagged total assets.

Change in Financial Liabilities (ChFL)

Based on Richardson et al. (2006), Change in Financial Liabilities is defined as

$$ChFL = \frac{FNL_t - FNL_{t-1}}{at_{t-1}}$$

where FNL_t are net financial liabilities defined in anomaly Change in Net Financial Assets and at_{t-1} are one-year lagged total assets.

Discretionary Accruals (DA)

Based on Dechow et al. (1995), Discretionary Accruals are residuals from following cross-sectional regression estimated for each two-digit SIC industry and year combination:

$$\frac{OA_{i,t}}{A_{i,t-1}} = \beta_0 \frac{1}{at_{i,t-1}} + \beta_1 \frac{(sale_{i,t} - sale_{i,t-1}) - ((rect_{i,t} - rect_{i,t-1})}{at_{i,t-1}} + \beta_2 \frac{ppegt_{i,t}}{at_{i,t-1}} + \epsilon_{i,t}$$

where

Growth in Inventory (GriI)

Based on Thomas and Zhang (2002), Growth in Inventor is defined as

$$GriI = \frac{invt_t - invt_{t-1}}{(at_t + at_{t-1})/2}$$

where $invt_t$ are inventories and at_t are total assets.

Inventory Change (ICh)

Based on Thomas and Zhang (2002), inventory change is defined as

$$ICh = \frac{invt_t - invt_{t-1}}{at_{t-1}}$$

where $invt_t$ are inventories and at_{t-1} are one-year lagged total assets.
Only firms with positive inventories in at least one year included in definition are included.

Inventory Growth (IGr)

Based on Belo and Lin (2011), inventory growth is defined as

$$IGr = \frac{invt_t - invt_{t-1}}{invt_{t-1}}$$

where $invt_t$ are inventories.

M/B and Accruals (MBaAC)

Based on Bartov and Kim (2004), M/B and Accruals is defined as

$$MBaAC = \begin{cases} 1 & \text{if stock is in low book-to-market } (BM_t) \text{ and high accrual } (Accr_t) \text{ quintiles} \\ -1 & \text{if stock is in high book-to-market } (BM_t) \text{ and low accrual } (Accr_t) \text{ quintiles} \\ 0 & \text{otherwise} \end{cases}$$

Accruals (Acc_t) are defined above, and book-to-market, book equity divided by market equity, (BM_t) is defined in category *Value*.

Net Working Capital Changes (NWCCh)

Based on Soliman (2008), net working capital changes are defined as

$$NWCCh = \frac{NWC_t - NWC_{t-1}}{at_{t-1}}$$

$NWC_t = (act_t - che_t) - (lct_t - dlc_t)$ is net working capital, where act_t are current assets, che_t is cash and cash equivalents, clt_t are current liabilities and dlc_t is debt in current liabilities.

Percent Operating Accruals (POA)

Based on Hafzalla et al. (2011), percent operating accruals are defined as

$$POA = \frac{ni_t - oancf_t}{|ni_t|}$$

where ni_t is net income and $oancf_t$ is cash flow from operations.

Percent Total Accruals (PTA)

Based on Hafzalla et al. (2011), percent total accruals are defined as

$$PTA = \frac{ni_t - (-sstk_t + prstkc_t + dv_t + oancf_t + ivncf_t + fincf_t)}{|ni_t|}$$

where ni_t is net income, $sstk_t$ sale of common and preferred stock, $prstkc_t$ is purchase of common and preferred stock, dv_t is total dividends, $oancf_t$ is cash flow from financing, $ivncf_t$ is cash flow from investment and $fincf_t$ is cash from from financing.

Total Accruals (TA)

Based on Richardson et al. (2006), total accruals are defined as

$$TA = \frac{TACCR_t - TACCR_{t-1}}{at_{t-1}}$$

where $TACCR_t = NCO_t + WC_t + NFNA_t$ NCO_t are net non-current operating assets defined in anomaly Change in Net Non-Current Operating Assets, WC_t is working capital defined in anomaly Change in Net Non-Cash Working Capital and $NFNA_t$ are net financial assets defined in anomaly Change in Net Financial Assets.

Intangibles

Asset Liquidity (AL)

Based on Ortiz-Molina and Phillips (2014), asset liquidity is defined as

$$AL = \frac{che_t + 0.75(act_t - che_t) + 0.5(at_t - act_t - gdwl_t - intan_t)}{at_{t-1}}$$

where at_{t-1} are one-year lagged total assets, act_t are current assets, che_t is cash and short-term investments, $gdwl_t$ is goodwill (0 if missing) and $intan_t$ are intangibles (0 if missing).

Asset Liquidity II (AL2)

Based on Ortiz-Molina and Phillips (2014), Asset Liquidity II is defined as

ΔSales - ΔAccounts Receivable (ChSChAR)

Based on Abarbanell and Bushee (1998), ΔSales - ΔAccounts Receivable is defined as

$$ChSChAR = \frac{sale_t - \frac{sale_{t-1}+sale_{t-2}}{2}}{\frac{sale_{t-1}+sale_{t-2}}{2}} - \frac{rect_t - \frac{rect_{t-1}+rect_{t-2}}{2}}{\frac{rect_{t-1}+rect_{t-2}}{2}}$$

where $sale_t$ is net sales and $rect_t$ are total receivables.
Only firms with positive two-year sales and two-year gross margin average are included.

ΔGross Margin - ΔSales (ChGMChS)

Based on Abarbanell and Bushee (1998), ΔGross Margin - ΔSales is defined as

$$ChSChAR = \frac{GM_t - \frac{GMs_{t-1}+GM_{t-2}}{2}}{\frac{GM_{t-1}+GM_{t-2}}{2}} - \frac{sale_t - \frac{sale_{t-1}+sale_{t-2}}{2}}{\frac{sale_{t-1}+sale_{t-2}}{2}}$$

where $sale_t$ is net sales and GM_t is gross margin, defined as $GM_t = sale_t - cogs_t$, where $cogs_t$ is cost of goods sold.
Only firms with positive two-year sales and two-year gross margin average are included.

Earnings Conservatism (EC)

Based on Francis et al. (2004),

$$EARN_it = \alpha_{i0} + \alpha_{i1} NEG_{it} + \beta_{i1} R_{it} + \beta_{i2} NEG_{it} R_{it} + e_{it}$$

in which $EARN_it = \frac{ib_t}{ME_t}$, where ib_t are earnings, ME_t is market equity defined in anomaly book-to-market in Section Value, R_{it} is i's stock 15-month return and NEG_{it} is defined as:

$$NEG_{it} = \begin{cases} 1 & \text{if } R_{it} < 0 \\ 0 & \text{otherwise} \end{cases}$$

Earnings Conservatism is defined as $EC = \frac{\beta_{i1}+\beta_{i2}}{\beta_{i1}}$

Earnings Persistence (EPe)

Based on Francis et al. (2004), Earnings Persistence is defined as the slope coefficient (beta) from the first-order autoregressive model using the ten-year rolling window for split-adjusted earnings per share. Split-adjusted earnings per share are defined as $EPS_t = \frac{epspx_t}{ajex_t}$.
Only firms with no missing required data over the ten-year rolling window are included.

Earnings Predictability (EPr)

Based on Francis et al. (2004), Earnings Predictability is defined as volatility of residuals from the first-order autoregressive model using the ten-year rolling window for split-adjusted earnings per share. Split-adjusted earnings per share are defined as $EPS_t = \frac{epspx_t}{ajex_t}$.
Only firms with no missing required data over the ten-year rolling window are included.

Earnings Timeliness (ET)

Based on Francis et al. (2004),

$$EARN_it = \alpha_{i0} + \alpha_{i1}NEG_{it} + \beta_{i1}R_{it} + \beta_{i2}NEG_{it}R_{it} + e_{it}$$

in which $EARN_it = \frac{ib_t}{ME_t}$, where ib_t are earnings, ME_t is market equity defined in anomaly book-to-market in Section Value, R_{it} is i's stock 15-month return and NEG_{it} is defined as:

$$NEG_{it} = \begin{cases} 1 & \text{if } R_{it} < 0 \\ 0 & \text{otherwise} \end{cases}$$

Earnings Timeliness is defined as R^2 from the regression.

Earning Smoothness (ES)

Based on Ortiz-Molina and Phillips (2014), earnings smoothness is defined as

$$ES = \frac{std(ELA_t)}{std(CFOA_t)}$$

where standard deviation is calculated over the ten-year rolling window and only firms with no missing required data over the ten-year history are included. Further

$$ELA_t = \frac{ib_t}{at_{t-1}}$$

and

$$CFOA_t = ib_t - (DCA_t - DCL_t - DCHE_t + DSTD_t - dp_t))$$

where ib_t are earnings and at_{t-1} is lagged total assets. DCA_t is one-year change in current assets, DCL_t is the one-year change in current liabilities, $DCHE_t$ is the one-year change in cash and short-term investments, $DSTD_t$ is the one-year change in debt in current liabilities and dp_t is depreciation and amortization.

Herfindahl Index (HI)

Based on Hou and Robinson (2006), Herfindahl index as a measure of industry concentration is defined as

$$HI = \frac{H_t + H_{t-1} + H_{t-2}}{3}$$

$H_t = \sum_{i=1}^{N_j} sale_{i,j}$, where $sale_{ij}$ is the sale of firm i in industry j and N_j is the total number of firms in the 3-digit SIC code defined industry.

Hiring rate (HR)

Based on Belo et al. (2014), hiring rate is defined as

$$HR = \frac{emp_{t-1} - empt - 2}{0.5emp_{t-1} + 0.5emp_{t-2}}$$

where emp_t is the number of employees. Stocks with $HR = 0$, often consequency of a stale information, are exluded.

Industry-adjusted Real Estate Ratio (IARER)

Based on Tuzel (2010),industry-adjusted real estate ratio is defined as

$$IARER = RER_t - \frac{\sum_{j=1}^{N_j} RER_{ij}}{N_j}$$

i.e. the real estate ratio minus its, 2-digit SIC code defined, industry average. Real estate ratio is defined as

$$RER_t = (fatb_t + fatl_t)/ppent_t$$

where $fatb_t$ is the sum of buildings at cost, $fatl_t$ is leases at cost and $ppegt_t$ is gross property, plant, and equipment.
Industries with less than five firms are excluded.

Industry-adjusted Organizational Capital-to-Assets (IaOCA)

Based on Eisfeldt and Papanikolaou (2013), Industry-adjusted Organizational Capital-to-Assets is defined as

$$IaOCA = \frac{OCA_t - \frac{\sum_{j=1}^{N_j} OCA_{ij}}{N_j}}{std(OCA_{ij})}$$

where $OCA_t = \frac{OC_t}{at_t}$ is organizational capital-to-assets, in which OC_t is organizational capital defined below in anomaly Org. Capital. Industry-adjusted organizational capital-to-assets is thus firm's org. capital industry demeaned and then divided by the standard deviation of org. capital within its industry.

Industry Concentration Assets (ICA)

Based on Hou and Robinson (2006), Industry Concentration Assets is Herfindahl index (HI), defined above, with total assets at_t as a measure of market share instead of sales $sale_t$.

Industry Concentration Book Equity (ICBE)

Based on Hou and Robinson (2006), Industry Concentration Book Equity is Herfindahl index (HI), defined above, with book equity BE_t defined in anomaly Book Equity / Market Equity.

Org. Capital (OC)

Based on Eisfeldt and Papanikolaou (2013), organizational capital is defined recursively. For the first year of stocks appearance in data, organizational capital is set equal to 4 times selling, general and administrative expense (0 if missing), i.e.

$$OC_{t_0} = 4 * xsga_{t_0}$$

All next years, organizational capital is defined as

$$OC_t = \frac{\frac{0.85 * OC_{t-1} + xsga_t}{cpi_t}}{at_t}$$

where cpi_t is and at_t are total assets.

R&D Capital-to-assets (RDCA)

Based on Li (2011), R&D Capital-to-assets is defined as

$$RDCA = \frac{xrd_t + 0.8xrd_{t-1} + 0.6xrd_{t-2} + 0.4xrd_{t-3} + 0.2xrd_{t-4}}{at_t}$$

where xrd_t are R&D expenses and at_t are total assets. Nominator is thus accumulated annual R&D expenses over the past five years with a linear depreciation rate of 20%.
Only firms with positive numerator and nonmissing xrd_t are included.

R&D Expenses-to-sales (RDES)

Based on Chan et al. (2001), R&D Expenses-to-sales is defined as

$$RDES = \frac{xrd_t}{sale_t}$$

where xrd_t is research and development expense and $sale_t$ are sales.
Only firms with positive xrd_t are included.

R&D / Market Value of Equity (RDM)

Based on Chan et al. (2001), R&D-to-market value of equity is defined as

$$RDM = \frac{xrd_t}{ME_t}$$

where xrd is research and development expense and $ME_t = prc_t * shrout_t$ is the market equity defined as price times shares outstanding, at the end of the previous year.

ΔSales - ΔInventory (SmI)

Based on Abarbanell and Bushee (1998), change in sales - change in inventory (ΔSales - ΔInventory) is defined as

$$SmI = \frac{sale_t - \frac{sale_{t-1}+sale_{t-2}}{2}}{\frac{sale_{t-1}+sale_{t-2}}{2}} - \frac{invt_t - \frac{invt_{t-1}+invt_{t-2}}{2}}{\frac{invt_{t-1}+invt_{t-2}}{2}}$$

where $sale_t$ is net sales and $invt_t$ is total inventories.
Annual rebalancing frequency.

Tangibility (TAN)

Based on Hahn and Lee (2009), tangibility is defined as

$$TAN = \frac{che_t + 0.715rect_t + 0.547invt_t + 0.535ppegt_t}{at_t}$$

where che_t are cash holdings, $rect_t$ are accounts receivable, $invt_t$ is inventory and $ppegt_t$ is property, plant and equipment.

Unexpected R&D Increases (URDI)

Based on Eberhart et al. (2004), unexpected R&D increases is binary variable defined as

$$URDI = \begin{cases} 1 & \text{if } (\frac{xrd_t}{revt_t} > 0.05) \text{ \& } (\frac{xrd_t}{at_t} > 0.05) \text{ \& } (\frac{xrd_t}{xrd_{t-1}} > 1.05) \text{ \& } (\frac{\frac{xrd_t}{at_t}}{\frac{xrd_{t-1}}{at_{t-1}}} > 1.05) \\ 0 & \text{otherwise} \end{cases}$$

where xrd_t are R&D expenditures, $revt_t$ is total revenue and at_t is total assets. $URDI = 1$, if revenue and R&D scaled by assets are greater than 5%, the yearly percentage change in R&D expenditures is greater than 5%; and R&D scaled by assets increased by more than 5%.

Whited-Wu Index (WWI)

Based on Whited and Wu (2006), Whited-Wu index is defined as

$$WWI_{it} = -0.091CF_t - 0.062DIVP_t + 0.021LDA_t - 0.044log(at_t) + 0.102ISG_t - 0.035(SG_t)$$

where

$$CF_T = \sqrt[4]{1 + \frac{\frac{ib_t+dp_t}{at_t}}{4}} - 1$$

where ib_t is income before extraordinary items, dp_t is depreciation and amortization, at_t are total assets, $DIVP_t$ is binary variable equal to one if firm pays cash dividends ($dvpsx_t > 0$) and 0 otherwise, $LDA_t = \frac{dltt_t}{at_t}$ is the long-term debt to total assets.

$$ISG_t = \frac{(\sum_{i=1}^{N_j} sale_{i,j})_t}{(\sum_{i=1}^{N_j} sale_{i,j})_t}$$

where $sale_{ij}$ is the sale of firm i in industry j and N_j is the total number of firms in the 3-digit SIC code defined industry including at least 3 firms.

$$SG_t = \sqrt[4]{1 + \frac{\frac{sale_t}{sale_{t-1}}}{4}} - 1$$

Investment

Asset Growth (AGr)

Based on Cooper et al. (2008), asset growth is defined as

$$AGr = \frac{at_t}{at_{t-1}}$$

where at_t are total assets.

Change Net Operating Assets (ChNOA)

Based on Hirshleifer et al. (2004), Change Net Operating Assets is defined as

$$ChNOA = \frac{NOA_t - NOA_{t-1}}{at_{t-1}}$$

where NOA_t are net operating assets defined below and $at_{t-1} arelaggedtotalassets$.

Changes in PPE and Inventory-to-Assets (ChPPEIA)

Based on Lyandres et al. (2007), Changes in PPE and Inventory-to-Assets is defined as

$$ChPPEIA_t = \frac{(ppegt_t - ppegt_{t-1}) + (invt_t - invt_{t-1})}{at_{t-1}}$$

wehere $ppegt_t$ is gross property, plant and equipment, $invt_t$ is total inventories and at_{t-1} are lagged total assets.

Composite Debt Issuance (CDI)

Based on Lyandres et al. (2007), Composite Debt Issuance is defined as

$$CDI = log(\frac{dltt_t + dlc_t}{dltt_{t-5} + dlc_{t-5}})$$

where $dltt_t$ is total long-term debt and dlc_t is debt in current liabilities.

ΔCAPEX - ΔIndustry CAPEX (CAPEX)

Based on Abarbanell and Bushee (1998), change in investment minus the change in industry investment (ΔCAPEX - ΔIndustry CAPEX) is defined as

$$CAPEX = \frac{capxv_t - \frac{capxv_t + capxv_{t-1}}{2}}{\frac{capxv_t + capxv_{t-1}}{2}} -$$

where capxv is capital expend property, plant and equipment.
Stocks in industries with less than 3 firms are excluded.

Debt Issuance (DI)

Based on Spiess and Affleck-Graves (1995), debt issuance is defined as

$$DI = \begin{cases} 1 & \text{if } dltis_t > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $dltis_t$ is long-term debt/issuance.

Growth in LTNOA (GriLTNOA)

Based on Fairfield et al. (2003), growth in long-term net operating assets is defined as

$$GriLTNOA = NOA_t - NOA_{t-1} - ACCR_t$$

, where NOA_t are net operating assets, defined below and $ACCR_t$ are accruals defined above in category *Accruals*.

Investment (INV)

Based on Titman et al. (2004), investment is defined as

$$INV = \frac{capx_t/revt_t}{avg_{3t}(\frac{capx}{revt})}$$

where $capx_t$ is capital expenditures, $revt_x$ is total revenue and $avg_{3t}()$ is average from the previous three years.
Stocks with revenue < \$10m are excluded.

Net Debt Finance (NDF)

Based on Bradshaw et al. (2006), Net Debt Finance is defined as

$$NDF_t = \frac{dltis_t - dltr_t + dlcch_t}{(at_t + at_{t-1})/2}$$

where $dltis_t$ is long-term debt issuance, $dltr_t$ is long-term debt reduction , $dlcch_t$ are current debt changes and at_t are total assets.

Net Equity Finance (NDF)

Based on Bradshaw et al. (2006), Net Equity Finance is defined as

$$NDF_t = \frac{sstk_t - prstkc_t - dv_t}{(at_t + at_{t-1})/2}$$

where $sstk_t$ is sale of common and preferred stock (0 if missing), $prstkc_t$ is purchase of common and preferred stock (0 if missing) , dv_t are cash dividend and at_t are total assets.

Net Operating Asset (NOA)

Based on Hirshleifer et al. (2004), net operating assets are defined as

$$NOA = \frac{OA_t - OL_t}{at_{t-1}}$$

OA_t and OL_t are operating assets and operating liabilities defined as $OA_t = at_t - che_t$ and $OL_t = at_t - dlc_t - dltt_t - mib_t - pstkrv_t - ceq_t$, where at_t is total assets, che_t is cash and short-term investment, dlc_t is current portion of long-term debt, $dltt$ is long-term debt, mib_t is minority interest, $pstkrv$ is preferred stock and ceq is common equity.

Noncurrent Operating Assets Changes (NOACh)

Based on Soliman (2008), noncurrent operating assets changes are defined as

$$NOACh = \frac{NCOA_t - NCOA_{t-1}}{at_t}$$

where $NCOA_t$ is noncurrent operating assets. Noncurrent operating assets are defined as

$$NCOA_t = (at_t - act_t - ivaeq_t) - (lt_t - lct_t - dltt_t)$$

, where at_t are total asssets, act_t are current assets, $ivaeq_t$ are investment and advances (0 if missing), lt_t are total liabilities, lct_t are current liabilities and $dltt_t$ is long-term debt.

Share Repurchases (SR)

Based on Ikenberry et al. (1995), share repurchases are defined as binary variable

$$SR = \begin{cases} 1 & \text{if } prstkc_t > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $prstkc_t$ is purchase of common and preferred stock.

Total XFIN (TXFIN)

Based on Bradshaw et al. (2006), total net external financing is defined as

$$TXFIN = \frac{sstk_t - dv_t - prstkc_t + dltis_t - dltr_t}{at_t}$$

where at_t are total assets, $sstk_t$ is sale of common and preferred stock (0 if missing), dv_t are cash dividends, $prstkc_t$ is purchase of common and preferred stock (0 if missing), $dltis_t$ is sale of long-term debt and $dltr_t$ is purchase of long-term debt.

Profitability

Asset Turnover (AT)

Based on Soliman (2008), asset turnover is defined as

$$AT = \frac{sale_t}{avg_{2t}(NOA)}$$

where NOA are net operating assets defined as $NOA = (at_t - che_t) - (lt_t - dltt_t - dlc_t - mib_t)$ and $avg_{2t}(NOA)$ is average NOA from two years. at_t are total assets, che_t is cash and cash equivalents, lt_t are total liabilities, $dltt_t$ is long-term debt, and dlc_t is debt in current liabilities and mib_t is minority interest (0 if missing). Firms with negative NOA and negative operating income ($oiadp$) are exluded.

Capital Turnover (CT)

Based on (Haugen and Baker, 1996), capital turnover is defined as

$$CT = \frac{sale_t}{at_{t-1}}$$

where $sale_t$ is sales and at_{t-1} are one-year lagged total assets.

Cash-based Operating Profitability (CBOP)

Based on Ball et al. (2016), cash-based operating profitability is defined as

$$CBOP = \big(revt_t - cogs_t - xsga_t + xrd_t - (rect_t - rect_{t-1}) - (invt_t - invt_{t-1}) -$$
$$-(xpp_t - xpp_{t-1}) + (drc_t + drlt_t - drc_t - drlt_t) + (rect_t - rect_{t-1}) + (ap_t - ap_{t-1})$$
$$+(xacc_t - xacc_{t-1})\big)/at_t$$

where at_t are total assets, $revt_t$ is total revenue, $cogs_t$ is cost of goods sold, $xsga_t$ are selling, general, and administrative expenses, xrd_t are research and development expenditures (0 if missing), $rect_t$ are accounts receivables, $invt_t$ is inventory, xpp_t are prepaid expenses, drc_t is current deferred revenue, $drlt_t$ is long-term deferred revenue, ap_t are accounts payable and $xacc_t$ are accrued expenses. Changes (in brackets) are all equal to 0 if missing.

Change in Asset Turnover (ChiAT)

Based on Soliman (2008), change in asset turnover is defined as

$$ChiAT = AT_t - AT_{t-1}$$

where AT_t is asset turnover defined above.

Earnings Consistency (EC)

Based on Alwathainani (2009), earnings consistency is defined as

$$EC = \sqrt[5]{\Pi_{i=1}^{5}(1 + eg_i)} - 1$$

where eg_i is earnings growth is defined as

$$eg_t = \frac{epspx_t - epspx_{t-1}}{\frac{|epsx_t| + |epsx_{t-1}|}{2}}$$

where $epspx_t$ are earnings per share excluding extraordinary items. Stocks with $|eg_t| > 6$ are deleted. Also stocks with the last two earnings growth with opposite signs are excluded ($eg_t * eg_{t-1}$)

Earnings / Price (EP)

Based on (Basu, 1977), earnings-to-price is defined as

$$EP = \frac{ib_t}{ME_t}$$

where ib_t is income before extraordinary items and $ME_t = prc_t * shrout_t$ is market equity, i.e. price times shares outstanding.
Firms with $ib_t \leq 0$ are excluded.

F-Score (FSc)

Based on Piotroski (2000), F-score is defined as the sum of nine binary variables (F1-F9) and is further limited only to firms in the highest quintile with respect to book-to-market

$$F = \sum_{i=1}^{9} F_i$$

Binary variables are defined as

$$\begin{aligned}
F1 = &\quad 1 \text{ if } ni_t > 0;\ 0 \text{ otherwise} \\
F2 = &\quad 1 \text{ if } oancf_t > 0 \ ; 0 \text{ otherwise} \\
F3 = &\quad 1 \text{ if } \tfrac{ni_t}{at_t} > \tfrac{ni_{t-1}}{at_{t-1}} \ ; 0 \text{ otherwise} \\
F4 = &\quad 1 \text{ if } oancf_t > ni_t \ ; 0 \text{ otherwise} \\
F5 = &\quad 1 \text{ if } \tfrac{dltt_t}{at_t} < \tfrac{dltt_{t-1}}{at_{t-1}} \ ; 0 \text{ otherwise} \\
F6 = &\quad 1 \text{ if } \tfrac{act_t}{lct_t} > \tfrac{act_{t-1}}{lct_{t-1}} \ ; 0 \text{ otherwise} \\
F7 = &\quad 1 if sstk_t - (pstk_t - pstk_{t-1}) \leq 0 \ ; 0 \text{ otherwise} \\
F8 = &\quad 1 \text{ if } \tfrac{oiadp_t}{sale_t} > \tfrac{oiadp_{t-1}}{sale_{t-1}} \ ; 0 \text{ otherwise} \\
F9 = &\quad 1 \text{ if } \tfrac{sale_t}{at_t} > \tfrac{sale_{t-1}}{at_{t-1}} \ ; 0 \text{ otherwise}
\end{aligned}$$

where ni_t is net income, $oancf_t$ is cash-flow from operating activities, at_t are total assets, $dltt_t$ is long term debt, act_t is current assets, lct_t are current liabilities, $sstk_t$ is sale of common and preferred stock, $pstk_t$ is , total preferred stock, $oiadp_t$ is operating income after depreciation and $sale_t$ is net sales.

Operating Profits to Assets (OPtA)

Based on Ball et al. (2016), operating profits to assets are defined as

$$OPtA = \frac{revt_t - cogs_t - xsga_t + xrd_t}{at_t}$$

where $revt_t$ is total revenue, $cogs_t$ is cost of goods sold, $xsga_t$ is SG&A, xrd_t are research and development expenditures and at_t are total assets.

O-Score (OSc)

Based on Dichev (1998), O-score is defined as

$$OSc = -1.32 - 0.4078log(\frac{at_t}{cpi_t}) + 6.03 * (\frac{dltt_t + dlc_t}{at_t}) - 1.43 * (\frac{act_t - lct_t}{at_t}) + 0.076 * (\frac{lct_t}{act_t}) -$$
$$-1.72 * (OENEG_t) - 2.37 * (\frac{ni_t}{at_t}) - 1.83 * (\frac{pi_t}{dp_t}) + 0.285 * (INTWO_t) - 0.521 * (\frac{ni_t - ni_{t-1}}{|ni_t| + |ni_{t-1}|})$$

where at_t are total assets, cpi_t is inflation, $dltt_t$ are long-term liabilities, dlc_t are short-term liabilities, act_t are current assets, lct_t are current liabilities, $OENEG_t$ is binary variable equal to one if $lt_t > at_t$ and 0 otherwise, ni_t is net income, $INTWO_t$ is binary variable equal to one if stock has negative net income in both previous years and 0 otherwise.
Only stocks with SIC codes from 1 to 3999 and from 5000 to 5999 are included.

Return on Net Operating Assets (RNOA)

Based on Soliman (2008), return on net operating assets is defined as

$$RNOA = \frac{oiadp_t}{NOA_{t-1}}$$

where NOA are net operating assets defined as $NOA_t = (at_t - che_t) - (lt_t - dltt_t - dlc_t - mib_t)$. at_t are total assets, che_t is cash and cash equivalents, lt_t are total liabilities, $dltt_t$ is long-term debt, and dlc_t is debt in current liabilities and mib_t is minority interest (0 if missing).
Firms with negative NOA and negative operating income ($oiadp$) are exluded.

Value

Assets-to-Market (AM)

Based on Fama and French (1992), assets-to-market is defined as

$$AM = \frac{at_t}{ME_t}$$

where at_t are assets total and ME_t is market equity.

Book Equity / Market Equity (BM)

Based on Fama and French (1992), book-to-market equity is defined as

$$BM = log(\frac{BE_t}{ME_t})$$

Market equity is price times shares outstanding, $ME_t = prc_t * shrout_t$. Book equity is defined conditional on missing items as

$$BE_t = seq_t - PS_t$$

where seq_t is total stockholders' equity, if missing then $seq_t = ceq_t + pstk_t$, or $seq_t = at_t - lt_t$, where ceq_t is tangible common equity, $pstk_t$ is preferred stock using liquidating value, at_t are total assets, lt_t are total liabilities,and PS_t is preferred stock measured using (ordered on availability) redemption, liquidating or par value, i.e. $pstkrv_t, pstkl_t, pstk_t$.

Cash Flow / Market Value of Equity (CM)

Based on Lakonishok et al. (1994), cash flow to market value of equity is defined as

$$CM = \frac{ib_t + dp_t}{ME_t}$$

where ib_t is net income, dp_t is depreciation and amortization and ME_t is market equity defined above in book-to-market equity anomaly.

Duration of Equity (DurE)

Based on Dechow et al. (2004), duration of equuity is defined as

$$DurE =$$

Enterprise Component of Book/Price (ECoBP)

Based on Penman et al. (2007), enterprise component of book/price is defined as

$$ECoBP = \frac{BE_t + ND_t}{ND_t + ME_t}$$

where BE_t and ME_t are book value of equity and market equiy, defined above in book-to-market equity anomaly. $ND_t = dltt_t + dlc_t + pstk_t + dvpa_t - tstkp_t - che_t$ is net debt, where che_t is cash and short-term investments, $dltt_t$ is long-term debt, dlc_t is debt in current liabilities, $pstk_t$ is preferred stock, $dvpa_t$ is preferred dividends in arrears and $tstkp_t$ is preferred treasury stock.

Enterprise Multiple (EM)

Based on Loughran and Wellman (2011), enterprise multiple is defined as

$$EM = \frac{EV_t}{oibdp_t}$$

where $oibdp_t$ is operating cash flow and EV_t is enterprise value defined as $EV_t = ME_t + dltt_t + dlc_t + pstk_t + dvpa_t - tstkp_t - che_t$. ME_t is market equity defined above in book-to-market equity anomaly, $dltt_t$ is long-term debt, dlc_t is debt in current liabilities, $pstk_t$ is preferred stock, $dvpa_t$ is preferred dividends in arrears, $tstkp_t$ is preferred treasury stock and che_t is cash and short-term investments.

Intangible Return (IR)

Based on Daniel and Titman (2006), intangible return is defined as residual from the following cross-sectional regression

$$log(r_{t-5,t}) = \beta_0 + \beta_1 BM_{t-5} + \beta_2 log(RB_{t-5,t}) + \epsilon_t$$

where $r_{t-5,t}$ is 5- year stock return, BM_{t-5} is 5-year-lagged book-to-market defined in anomaly Book Equity / Market Equity and $RB_{t-5,t} = log(\frac{BE_t}{BE_{t-5}) - \sum_{p=t-5}^{t-1}(r_p - log(\frac{P_p}{P_{p-1}}))}$ in which BE_t is the book equity defined in anomaly Book Equity / Market Equity , r_p is the stockr return for year p and P_p is the price at the end of year p.

Leverage (Lvrg)

Based on Bhandari (1988), leverage is defined as

$$Lvrg = \frac{dltt_t + dlc_t}{ME_t}$$

where $dltt_t$ is long-term debt, dlc_t is debt in current liabilities and $ME_t = prc_t * shrout_t$ is market equity defined in anomaly of earnings/price.

Leverage Component of Book/Price (LCoBP)

Based on Penman et al. (2007), leverage component of book/price is defined as

$$LCoBP = BE_t - ECoBP_t$$

where BE_t is book value of equity defined above in book-to-market equity anomaly, and $ECoBP_t$ is enterprise component of book/price defined above.

Net Payout Yield (NPY)

Based on Boudoukh et al. (2007), net payout yield is defined as

$$NPY = \frac{dvc_t + prstkc_t - sstk_t}{ME_t}$$

where dvc_t are dividends common/ordinary, $prstkc_t$ is purchase of common and preferred stock, $sstk_t$ is sale of common and preferred stock and ME_t is market equity.

Operating Leverage (OL)

Based on Novy-Marx (2010), operating leverage is defined

$$OL = \frac{xsga_t + cogs_t}{at_t}$$

where $xsga_t$ is SG&A, $cogs_t$ is cost of goods sold and at_t are total assets.

Payout Yield (PY)

Based on Boudoukh et al. (2007), payout yield is defined as

$$PY = \frac{dvc_t + prstkc_t - (pstkrv_t + pstkrv_{t-1})}{ME_t}$$

where dvc_t are dividends common/ordinary, $prstkc_t$ is purchase of common and preferred stock, $pstkrv_t$ is preferred stock/redemption and ME_t is market equity.

Sales Growth (SaGr)

Based on Lakonishok et al. (1994), sales growth is defined as

$$SaGr = \frac{5SGR_t + 4SGR_{t-1} + 3SGR_{t-2} + 2SGR_{t-3} + 1SGR_{t-4}}{15}$$

where SGR_t is the rank of firm in year t based on the simple sales growth defined as $SG = sale_t/sale_{t-1}$.

Sustainable Growth (SuGr)

Based on Lockwood and Prombutr (2010), sustainable growth is defined as $SuGr = BE_t/BE_{t-1}$, where BE_t is book equity defined above in book-to-market equity anomaly.

Sales/Price (SP)

Based on Barbee Jr et al. (1996), sales-to-price is defined as $SP = revt_t/ME_t$, where $revt_t$ is total revenue and ME_t is the market equity defined above in book-to-market equity anomaly.

Frictions

I/B/E/S

References

Abarbanell, Jeffery S, and Brian J Bushee, 1998, Abnormal returns to a fundamental analysis strategy, *Accounting Review* 19–45.

Acharya, Viral V, and Lasse Heje Pedersen, 2005, Asset pricing with liquidity risk, *Journal of financial Economics* 77, 375–410.

Alwathainani, Abdulaziz M, 2009, Consistency of firms' past financial performance measures and future returns, *The British Accounting Review* 41, 184–196.

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of financial markets* 5, 31–56.

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of financial Economics* 17, 223–249.

Anderson, Evan W, Eric Ghysels, and Jennifer L Juergens, 2005, Do heterogeneous beliefs matter for asset pricing?, *The Review of Financial Studies* 18, 875–924.

Ang, Andrew, Joseph Chen, and Yuhang Xing, 2006a, Downside risk, *The Review of Financial Studies* 19, 1191–1239.

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

Asparouhova, Elena, Hendrik Bessembinder, and Ivalina Kalcheva, 2010, Liquidity biases in asset pricing tests, *Journal of Financial Economics* 96, 215–237.

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.

Bali, Turan G, Lin Peng, Yannan Shen, and Yi Tang, 2013, Liquidity shocks and stock market reactions, *The Review of Financial Studies* 27, 1434–1485.

Ball, Ray, Joseph Gerakos, Juhani T Linnainmaa, and Valeri Nikolaev, 2016, Accruals, cash flows, and operating profitability in the cross section of stock returns, *Journal of Financial Economics* 121, 28–45.

Banz, Rolf W, 1981, The relationship between return and market value of common stocks, *Journal of financial economics* 9, 3–18.

Barbee Jr, William C, Sandip Mukherji, and Gary A Raines, 1996, Do sales–price and debt–equity explain stock returns better than book–market and firm size?, *Financial Analysts Journal* 52, 56–60.

Barber, Brad, Reuven Lehavy, Maureen McNichols, and Brett Trueman, 2001, Can investors profit from the prophets? security analyst recommendations and stock returns, *The Journal of Finance* 56, 531–563.

Barry, Christopher B, and Stephen J Brown, 1984, Differential information and the small firm effect, *Journal of Financial Economics* 13, 283–294.

Barth, Mary E, and Amy P Hutton, 2004, Analyst earnings forecast revisions and the pricing of accruals, *Review of accounting studies* 9, 59–96.

Bartov, Eli, and Myungsun Kim, 2004, Risk, mispricing, and value investing, *Review of Quantitative Finance and Accounting* 23, 353–376.

Basu, Sanjoy, 1977, Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis, *The journal of Finance* 32, 663–682.

Belo, Frederico, and Xiaoji Lin, 2011, The inventory growth spread, *The Review of Financial Studies* 25, 278–313.

Belo, Frederico, Xiaoji Lin, and Santiago Bazdresch, 2014, Labor hiring, investment, and stock return predictability in the cross section, *Journal of Political Economy* 122, 129–177.

Bhandari, Laxmi Chand, 1988, Debt/equity ratio and expected common stock returns: Empirical evidence, *The journal of finance* 43, 507–528.

Blitz, David, Joop Huij, and Martin Martens, 2011, Residual momentum, *Journal of Empirical Finance* 18, 506–521.

Blume, Marshall E, and Frank Husic, 1973, Price, beta, and exchange listing, *The Journal of Finance* 28, 283–299.

Bondt, Werner FM, and Richard Thaler, 1985, Does the stock market overreact?, *The Journal of finance* 40, 793–805.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R Roberts, 2007, On the importance of measuring payout yield: Implications for empirical asset pricing, *The Journal of Finance* 62, 877–915.

Bradshaw, Mark T, Scott A Richardson, and Richard G Sloan, 2006, The relation between corporate financing activities, analysts' forecasts and stock returns, *Journal of Accounting and Economics* 42, 53–85.

Chan, Louis KC, Josef Lakonishok, and Theodore Sougiannis, 2001, The stock market valuation of research and development expenditures, *The Journal of Finance* 56, 2431–2456.

Chen, Tianqi, and Tong He, 2017, xgboost: extreme gradient boosting .

Chordia, Tarun, Avanidhar Subrahmanyam, and V Ravi Anshuman, 2001, Trading activity and expected stock returns, *Journal of Financial Economics* 59, 3–32.

Cooper, Michael J, Huseyin Gulen, and Michael J Schill, 2008, Asset growth and the cross-section of stock returns, *The Journal of Finance* 63, 1609–1651.

Da, Zhi, and Mitch Warachka, 2011, The disparity between long-term and short-term forecasted earnings growth, *Journal of Financial Economics* 100, 424–442.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *The Journal of Finance* 61, 1605–1643.

Datar, Vinay T, Narayan Y Naik, and Robert Radcliffe, 1998, Liquidity and stock returns: An alternative test, *Journal of Financial Markets* 1, 203–219.

Dechow, Patricia M, Richard G Sloan, and Mark T Soliman, 2004, Implied equity duration: A new measure of equity risk, *Review of Accounting Studies* 9, 197–228.

Dechow, Patricia M, Richard G Sloan, and Amy P Sweeney, 1995, Detecting earnings management, *Accounting review* 193–225.

Dichev, Ilia D, 1998, Is the risk of bankruptcy a systematic risk?, *the Journal of Finance* 53, 1131–1147.

Diebold, Francis X, and Roberto S Mariano, 1995, Comparing predictive accuracy, *Journal of Business & Economic Statistics* .

Diether, Karl B, Christopher J Malloy, and Anna Scherbina, 2002, Differences of opinion and the cross section of stock returns, *The Journal of Finance* 57, 2113–2141.

Driscoll, John C, and Aart C Kraay, 1998, Consistent covariance matrix estimation with spatially dependent panel data, *The Review of Economics and Statistics* 80, 549–560.

Eberhart, Allan C, William F Maxwell, and Akhtar R Siddique, 2004, An examination of long-term abnormal stock returns and operating performance following r&d increases, *The Journal of Finance* 59, 623–650.

Eisfeldt, Andrea L, and Dimitris Papanikolaou, 2013, Organization capital and the cross-section of expected returns, *The Journal of Finance* 68, 1365–1406.

Elgers, Pieter T, May H Lo, and Ray J Pfeiffer Jr, 2001, Delayed security price adjustments to financial analysts' forecasts of annual earnings, *The Accounting Review* 76, 613–632.

Fairfield, Patricia M, J Scott Whisenant, and Teri Lombardi Yohn, 2003, Accrued earnings and growth: Implications for future profitability and market mispricing, *The accounting review* 78, 353–371.

Fama, Eugene F, and Kenneth R French, 1992, The cross-section of expected stock returns, *the Journal of Finance* 47, 427–465.

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 Kenneth R French, 2017, International tests of a five-factor asset pricing model, *Journal of Financial Economics* 123, 441–463.

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

Fong, Kingsley, Craig Holden, and Ondrej Tobek, 2017, Are volatility over volume liquidity proxies useful for global or us research? .

Francis, Jennifer, Ryan LaFond, Per M Olsson, and Katherine Schipper, 2004, Costs of equity and earnings attributes, *The accounting review* 79, 967–1010.

Frankel, Richard, and Charles MC Lee, 1998, Accounting valuation, market expectation, and cross-sectional stock returns, *Journal of Accounting and economics* 25, 283–319.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber, 2017, Dissecting characteristics nonparametrically, Technical report, National Bureau of Economic Research.

Friedman, Jerome H, 2001, Greedy function approximation: a gradient boosting machine, *Annals of statistics* 1189–1232.

George, Thomas J, and Chuan-Yang Hwang, 2004, The 52-week high and momentum investing, *The Journal of Finance* 59, 2145–2176.

Green, Jeremiah, John RM Hand, and X Frank Zhang, 2017, The characteristics that provide independent information about average us monthly stock returns, *The Review of Financial Studies* hhx019.

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

Gu, Shihao, Bryan T Kelly, and Dacheng Xiu, 2018, Empirical asset pricing via machine learning .

Hafzalla, Nader, Russell Lundholm, and E Matthew Van Winkle, 2011, Percent accruals, *The Accounting Review* 86, 209–236.

Hahn, Jaehoon, and Hangyong Lee, 2009, Financial constraints, debt capacity, and the cross-section of stock returns, *The Journal of Finance* 64, 891–921.

Harvey, Campbell R, Yan Liu, and Heqing Zhu, 2016, ... and the cross-section of expected returns, *The Review of Financial Studies* 29, 5–68.

Harvey, Campbell R, and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *The Journal of Finance* 55, 1263–1295.

Haugen, Robert A, and Nardin L Baker, 1996, Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401–439.

Hawkins, Eugene H, Stanley C Chamberlin, and Wayne E Daniel, 1984, Earnings expectations and security prices, *Financial Analysts Journal* 40, 24–38.

Heston, Steven L, and Ronnie Sadka, 2008, Seasonality in the cross-section of stock returns, *Journal of Financial Economics* 87, 418–445.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets?, *Journal of Accounting and Economics* 38, 297–331.

Hou, Kewei, and David T Robinson, 2006, Industry concentration and average stock returns, *The Journal of Finance* 61, 1927–1956.

Hou, Kewei, Chen Xue, and Lu Zhang, 2017, Replicating anomalies, Technical report, National Bureau of Economic Research.

Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of financial economics* 39, 181–208.

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, and Sebastian Müller, 2017a, ... and nothing else matters? on the dimensionality and predictability of international stock returns .

Jacobs, Heiko, and Sebastian Müller, 2017b, Anomalies across the globe: Once public, no longer existent? .

Jacobs, Heiko, and Sebastian Müller, 2017c, Measuring mispricing .

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *The Journal of finance* 45, 881–898.

Jegadeesh, Narasimhan, Joonghyuk Kim, Susan D Krische, and Charles Lee, 2004, Analyzing the analysts: When do recommendations add value?, *The journal of finance* 59, 1083–1124.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *The Journal of finance* 48, 65–91.

Kelly, Bryan, and Hao Jiang, 2014, Tail risk and asset prices, *The Review of Financial Studies* 27, 2841–2871.

Kot, Hung Wan, and Kalok Chan, 2006, Can contrarian strategies improve momentum profits .

Kyle, Albert S, and Anna A Obizhaeva, 2016, Market microstructure invariance: Empirical hypotheses, *Econometrica* 84, 1345–1404.

La Porta, Rafael, 1996, Expectations and the cross-section of stock returns, *The Journal of Finance* 51, 1715–1742.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny, 1994, Contrarian investment, extrapolation, and risk, *The journal of finance* 49, 1541–1578.

Lee, Charles, and Bhaskaran Swaminathan, 2000, Price momentum and trading volume, *the Journal of Finance* 55, 2017–2069.

Lee, Kuan-Hui, 2011, The world price of liquidity risk, *Journal of Financial Economics* 99, 136–161.

Lewellen, Jonathan, et al., 2015, The cross-section of expected stock returns, *Critical Finance Review* 4, 1–44.

Li, Dongmei, 2011, Financial constraints, r&d investment, and stock returns, *The Review of Financial Studies* 24, 2974–3007.

Lockwood, Larry, and Wikrom Prombutr, 2010, Sustainable growth and stock returns, *Journal of Financial Research* 33, 519–538.

Loughran, Tim, and Jay W Wellman, 2011, New evidence on the relation between the enterprise multiple and average stock returns, *Journal of Financial and Quantitative Analysis* 46, 1629–1650.

Lyandres, Evgeny, Le Sun, and Lu Zhang, 2007, The new issues puzzle: Testing the investment-based explanation, *The Review of Financial Studies* 21, 2825–2855.

McLean, R David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *The Journal of Finance* 71, 5–32.

Moskowitz, Tobias J, and Mark Grinblatt, 1999, Do industries explain momentum?, *The Journal of Finance* 54, 1249–1290.

Newey, Whitney K, and Kenneth D West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica: Journal of the Econometric Society* 703–708.

Novy-Marx, Robert, 2010, Operating leverage, *Review of Finance* 15, 103–134.

Novy-Marx, Robert, 2012, Is momentum really momentum?, *Journal of Financial Economics* 103, 429–453.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Ortiz-Molina, Hernán, and Gordon M Phillips, 2014, Real asset illiquidity and the cost of capital, *Journal of Financial and Quantitative Analysis* 49, 1–32.

Palazzo, Berardino, 2012, Cash holdings, risk, and expected returns, *Journal of Financial Economics* 104, 162–185.

Pástor, L'uboš, and Robert F Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political economy* 111, 642–685.

Penman, Stephen H, Scott A Richardson, and Irem Tuna, 2007, The book-to-price effect in stock returns: accounting for leverage, *Journal of Accounting Research* 45, 427–467.

Piotroski, Joseph D, 2000, Value investing: The use of historical financial statement information to separate winners from losers, *Journal of Accounting Research* 1–41.

Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share issuance and cross-sectional returns, *The Journal of Finance* 63, 921–945.

Richardson, Scott A, Richard G Sloan, Mark T Soliman, and Irem Tuna, 2006, The implications of accounting distortions and growth for accruals and profitability, *The Accounting Review* 81, 713–743.

Sloan, Allan, 1996, Create an account or log in, *Accounting review* 71, 289–315.

Soliman, Mark T, 2008, The use of dupont analysis by market participants, *The Accounting Review* 83, 823–853.

Spiess, D Katherine, and John Affleck-Graves, 1995, Underperformance in long-run stock returns following seasoned equity offerings, *Journal of Financial Economics* 38, 243–267.

Thomas, Jacob K, and Huai Zhang, 2002, Inventory changes and future returns, *Review of Accounting Studies* 7, 163–187.

Titman, Sheridan, KC John Wei, and Feixue Xie, 2004, Capital investments and stock returns, *Journal of financial and Quantitative Analysis* 39, 677–700.

Tobek, Ondrej, and Martin Hronec, 2018, Does source of fundamental data matter? .

Tuzel, Selale, 2010, Corporate real estate holdings and the cross-section of stock returns, *The Review of Financial Studies* 23, 2268–2302.

Whited, Toni M, and Guojun Wu, 2006, Financial constraints risk, *The Review of Financial Studies* 19, 531–559.

Zhang, X, 2006, Information uncertainty and stock returns, *The Journal of Finance* 61, 105–137.

Online Appendix

Table XVIII
Industries in Datastream level 3 classification and corresponding four digit SIC

Datastream lvl 3 industry	SIC codes
Automobiles & Parts	3011, 3510, 3714, 3751, 5013
Basic Resources	800, 1000, 1040, 1090, 1220, 1221, 2421, 2600, 2611, 2621, 2631, 3310, 3312, 3317, 3330, 3334, 3350, 3360, 3444, 3460, 3720, 5050, 5051
Chemicals	2810, 2820, 2821, 2833, 2851, 2860, 2870, 2890, 2891, 2990, 3080, 3081, 3341, 5160
Construct. & Material	1400, 1540, 1600, 1623, 1731, 2400, 2430, 2950, 3211, 3231, 3241, 3250, 3270, 3272, 3281, 3290, 3430, 3440, 3442, 3448, 5031, 5070, 5072
Financial Services(3)	6111, 6141, 6153, 6159, 6162, 6163, 6172, 6189, 6200, 6211, 6221, 6282, 6361, 6500, 6510, 6770, 6795, 6798, 6799, 8880, 8888, 9995
Food & Beverage	100, 200, 900, 2000, 2011, 2013, 2015, 2020, 2024, 2030, 2033, 2040, 2050, 2052, 2060, 2070, 2080, 2082, 2086, 2090, 2092
Healthcare	2590, 2800, 2834, 2835, 2836, 3060, 3821, 3826, 3841, 3842, 3843, 3844, 3845, 3851, 4100, 5047, 6324, 8000, 8011, 8050, 8051, 8060, 8062, 8071, 8082, 8090, 8093, 8300, 8731
Ind. Goods & Services	1700, 2390, 2650, 2670, 2673, 2750, 2761, 3050, 3086, 3089, 3221, 3320, 3357, 3390, 3411, 3412, 3443, 3451, 3452, 3470, 3480, 3490, 3523, 3524, 3530, 3531, 3532, 3537, 3540, 3541, 3550, 3555, 3560, 3561, 3562, 3564, 3567, 3569, 3575, 3580, 3585, 3590, 3600, 3612, 3613, 3620, 3621, 3634, 3640, 3669, 3670, 3672, 3677, 3678, 3679, 3690, 3711, 3713, 3715, 3721, 3724, 3728, 3730, 3743, 3760, 3812, 3822, 3823, 3824, 3825, 3827, 3829, 3861, 3910, 4011, 4013, 4210, 4213, 4231, 4400, 4412, 4513, 4700, 4731, 4950, 4953, 4955, 4961, 5000, 5063, 5065, 5080, 5082, 5084, 5090, 5099, 6099, 6794, 7320, 7350, 7359, 7361, 7363, 7374, 7377, 7380, 7381, 7384, 7385, 7389, 7829, 8111, 8200, 8351, 8600, 8700, 8711, 8734, 8741, 8742, 8744, 9721
Insurance	6311, 6321, 6331, 6351, 6411
Media	2711, 2721, 2731, 2732, 2741, 2780, 4832, 4833, 4841, 7310, 7311, 7330, 7331, 7819, 7822, 8900
Oil & Gas	1311, 1381, 1382, 1389, 2911, 3533, 4522, 4610, 4900, 5171, 5172, 6792
Pers & Househld Goods	1531, 2100, 2111, 2200, 2211, 2221, 2250, 2253, 2273, 2300, 2320, 2330, 2340, 2451, 2452, 2510, 2511, 2520, 2522, 2531, 2540, 2771, 2840, 2842, 2844, 3021, 3100, 3220, 3260, 3420, 3433, 3630, 3651, 3716, 3790, 3873, 3911, 3931, 3942, 3944, 3949, 3950, 3960, 5020, 5030, 5064, 5130, 5150, 5190, 6552
Real Estate	6519, 6531
Retail	700, 2790, 3140, 4220, 5094, 5010, 5110, 5122, 5140, 5141, 5180, 5200, 5211, 5271, 5311, 5331, 5399, 5400, 5411, 5412, 5500, 5531, 5600, 5621, 5651, 5661, 5700, 5712, 5731, 5734, 5735, 5912, 5940, 5944, 5945, 5960, 5961, 5990, 6399, 7200, 7340, 7500, 7600, 7841
Technology	3559, 3570, 3571, 3572, 3576, 3577, 3578, 3579, 3661, 3663, 3674, 3695, 4899, 5040, 5045, 7370, 7371, 7372, 7373
Telecommunications	4812, 4813, 4822
Travel & Leisure	1520, 3652, 3990, 4512, 4581, 5810, 5812, 6512, 6513, 6532, 7000, 7011, 7510, 7812, 7830, 7900, 7948, 7990, 7997
Utilities	4911, 4922, 4923, 4924, 4931, 4932, 4941, 4991, 5900
Banks	6021, 6022, 6029, 6035, 6036, 6199

Table XIX
Performance of shrinkage strategies estimated on stocks outside US: WLS regressions

The table shows returns of a shrinkage strategy as described in Table VI that is estimated on individual stocks from the US, US & Japan, US & Europe, or US & Japan & Europe. We adjust the performance of the mixing strategy for five Fama-French factors from individual regions. Standard errors in t-statistics are HAC adjusted, as in Newey and West (1987) with 12 lags. The returns are in percentage points per month.

	Equal-weighted				Value-weighted			
	USA	Europe	Japan	Asia Pacific	USA	Europe	Japan	Asia Pacific
Evidence from the US								
Mean Return	1.620	1.600	1.240	1.680	0.968	1.200	1.020	1.360
	(3.580)	(4.980)	(4.690)	(5.680)	(2.520)	(4.120)	(3.360)	(3.970)
FF5 alpha	1.060	0.875	1.020	1.400	0.477	0.571	0.823	0.941
	(3.390)	(4.070)	(4.220)	(4.770)	(1.650)	(2.290)	(2.790)	(2.470)
Evidence from the US & Japan								
Mean Return	1.550	1.360	1.360	1.460	0.798	1.020	0.955	1.010
	(3.710)	(4.570)	(4.900)	(5.330)	(2.120)	(3.740)	(2.730)	(3.130)
Diff wrt the US	-0.072	-0.233	0.116	-0.216	-0.170	-0.174	-0.068	-0.347
	(-0.588)	(-1.980)	(0.896)	(-1.490)	(-1.130)	(-1.270)	(-0.314)	(-1.970)
FF5 alpha	0.995	0.734	1.140	1.150	0.255	0.436	0.765	0.577
	(4.000)	(3.380)	(4.670)	(4.100)	(0.864)	(1.880)	(2.240)	(1.540)
Evidence from the US & Europe								
Mean Return	1.530	1.810	1.280	2.000	1.050	1.270	1.010	1.390
	(3.230)	(5.230)	(4.880)	(6.000)	(2.290)	(4.140)	(3.150)	(3.670)
Diff wrt the US	-0.088	0.213	0.038	0.319	0.085	0.071	-0.016	0.025
	(-1.290)	(2.600)	(0.490)	(3.220)	(0.487)	(0.615)	(-0.108)	(0.130)
FF5 alpha	0.972	0.995	1.070	1.540	0.458	0.531	0.747	0.864
	(3.170)	(4.450)	(4.560)	(5.260)	(1.590)	(2.230)	(2.420)	(2.230)
Evidence from the US & Japan & Europe								
Mean Return	1.560	1.760	1.350	2.080	0.934	1.070	0.917	1.430
	(3.310)	(5.170)	(5.170)	(7.120)	(2.030)	(3.570)	(2.870)	(3.640)
Diff wrt the US	-0.063	0.168	0.108	0.406	-0.034	-0.122	-0.106	0.067
	(-0.711)	(2.050)	(1.090)	(3.550)	(-0.179)	(-0.782)	(-0.599)	(0.262)
FF5 alpha	1.010	0.952	1.150	1.620	0.320	0.320	0.633	0.699
	(3.460)	(4.490)	(5.130)	(6.620)	(1.080)	(1.430)	(2.110)	(1.910)