An EDHEC-Risk Institute Publication

Smart Beta Replication Costs

February 2017

with the support of

Amundi ETF INDEXING SMART BETA
ASSET MANAGEMENT

EDHEC-RISK Institute

Printed in France, February 2017. Copyright EDHEC 2017.
Document prepared as part of the Amundi ETF, Indexing and Smart Beta "ETF and Passive Investment Strategies" research chair at EDHEC-Risk Institute.
The opinions expressed in this study are those of the authors and do not necessarily reflect those of EDHEC Business School.

Table of Contents

Executive Summary 5

Introduction 13

1. Methodology 17

2. Descriptive Statistics on Transaction Costs 27

3. Analysing Smart Beta Strategies 41

Conclusions 61

References 67

About Amundi ETF, Indexing & Smart Beta 73

About EDHEC-Risk Institute 75

EDHEC-Risk Institute Publications and Position Papers (2014-2017) 79

About the Authors

Mikheil Esakia is a Quantitative Research Analyst at ERI Scientific Beta. He holds a Master of Science degree in Finance from EDHEC Business School (with distinction) and a Bachelor's degree in Business Administration from the Free University of Tbilisi.

Felix Goltz is Head of Applied Research at EDHEC-Risk Institute. He carries out research in empirical finance and asset allocation, with a focus on alternative investments and indexing strategies. His work has appeared in various international academic and practitioner journals and handbooks. He obtained a PhD in finance from the University of Nice Sophia-Antipolis after studying economics and business administration at the University of Bayreuth and EDHEC Business School.

Sivagaminathan Sivasubramanian is a Quantitative Research Analyst at ERI Scientific Beta. He holds a Master of Science degree in Financial Markets from EDHEC Business School as well as a first-class Bachelor's degree (with distinction) in Computer Science and Engineering. Prior to joining ERI Scientific Beta, he worked as a software programmer for a number of years.

Jakub Ulahel is a Quantitative Equity Analyst with ERI Scientific Beta. He holds an MSc in Financial Markets from EDHEC Business School in France, where he was included in the Dean's list. His work focuses on the analysis of alternative equity portfolios. Previously, he worked as Business Analyst for Microsoft after graduating from the Charles University in Prague with a bachelor´s degree in Economics.

Executive Summary

Executive Summary

Introduction

An important issue with smart beta strategies is that they typically entail higher replication costs than cap-weighted market indices. While this is obviously true, the crux of the question is not whether transaction costs are higher but whether, after accounting for such costs, there are any benefits in terms of net returns. A reasonable expectation from an investor's perspective is that providers should disclose the estimated level of transaction costs generated by their strategies so as to allow for information on net returns. However, providers typically fail to make explicit adjustments for transaction costs and satisfy themselves by reporting gross returns, leaving it to other market participants to figure out what exactly the transaction costs amount to. This study sets out to apply methods for explicit cost measurement and to thus draw conclusions on smart beta strategies.

Transaction Cost Estimates

Easily accessible transaction cost estimates

A first important objective of this study is to test methods which provide easy access to direct transaction cost estimates.

Transaction cost estimates for smart beta strategies are hard to obtain in practice because an accurate estimation in principle requires intraday high frequency data. One needs to observe trades and quotes within the trading day to come up with cost measures. However, not only is such data difficult to access, it is also difficult to use. The increasing frequency of trading has led to a huge amount of tick by tick price data that requires massive computational power for analysis, with some researchers arguing that the growth of high frequency equity even outpaces the growth of computing power. Moreover, tick data requires matching procedures for prices and quotes so that the quality of databases and the cleaning procedures becomes a prime concern. Moreover, high frequency data only covers relatively short time periods, making it impossible to evaluate long-term track records of smart beta strategies.

Recent research has shown that there are effective ways of estimating transaction cost variables that are only observable at high frequency, based on lower frequency (daily) data. We draw on recent advances in microstructure research to extract measures of transaction costs from daily data, such as the daily range between high and low prices and the closing bid-ask spread. Using daily data allows us to analyse longer time periods than would be possible if drawing on high frequency data. Moreover, the methods we use are not computationally intensive and they draw on easily available data, making them easily replicable for practitioners who wish to analyse smart beta strategies.

We follow two types of spread estimation methods based on daily data – one based on Corwin and Schultz (2012) who use daily range measures such as high and low prices to estimate daily spreads (hereinafter referred to as the *range-based* spread estimator), and the other based on Chung and Zhang (2014) who use daily closing quoted bid and ask prices to estimate daily spreads (hereinafter referred to as the closing spread estimator).

While there is substantial literature suggesting that such measures are highly correlated with high frequency cost measures, our assessment indeed confirms that low frequency measures reliably capture the level of costs. In particular, we show that our measures capture information

Executive Summary

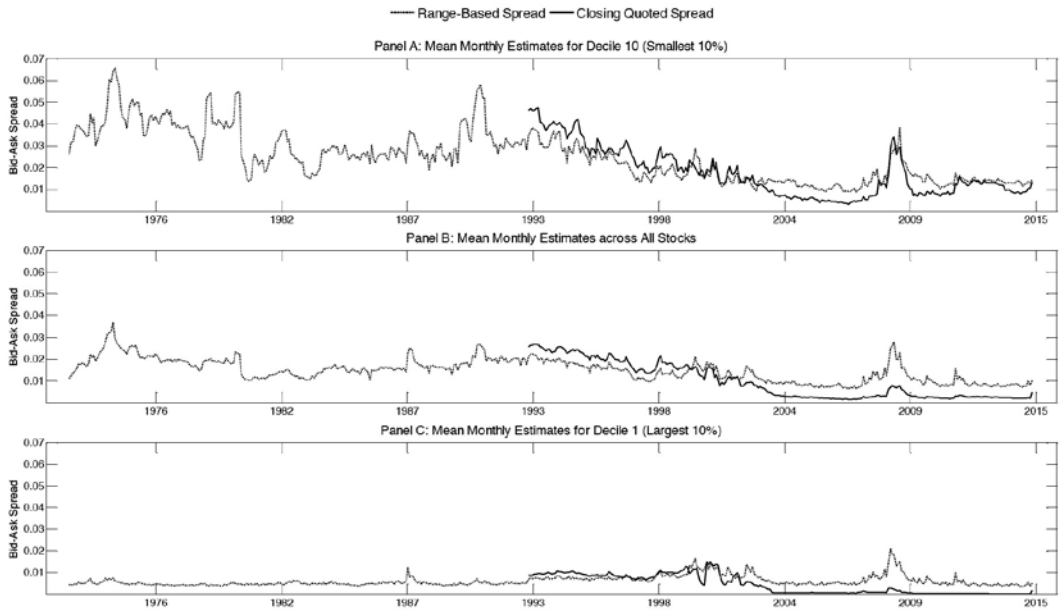
content of transaction costs (effective spreads) reported by trading venues in compliance with Rule 605 regulations. They also align well with effective spreads extracted from high frequency trade and quote data (TAQ). Compared to estimates from high frequency data, our cost measures are, however, somewhat conservative in that they tend to slightly overestimate cost levels. This means that any conclusions about the viability of smart beta strategies in the face of transaction costs will also tend to be on the conservative side.

While we apply our cost estimates to a range of smart beta strategies to draw conclusions about cost levels, it is worth emphasising that our transaction cost measurement approach can easily be applied to testing additional strategies. Using methods such as those in this paper could help the industry make cost estimates more widely available given the computational ease and widely accessible data such cost estimates are based on.

Transaction cost levels across stocks and over time

The following exhibit shows results for the average spread across all stocks, as well as the average spreads for the largest and smallest stocks in our universe. Large and small stocks are taken as the top and bottom deciles every year by market capitalisation (as of the last trading day of the previous year). The 3,000 stocks available in every quarter of a given year are aggregated for the decile selection. The number of unique stocks may thus be greater than 3,000 in a given year. Monthly average spread estimates are then calculated for these deciles.

Effective Spread Estimates: Top 3,000 US stock universe
The figure shows the mean monthly spread estimates based on two estimators – the Range-Based Spread Estimator and the Closing Quoted Spread Estimator. Reported spreads are mean monthly 2-way spread estimates. Our sample universe consists of the 3,000 largest ordinary common stocks in the United States in each quarter based on market capitalisation. As the universe is re-sampled every quarter there may be more than 3,000 stocks in a given year. The daily spread estimate of each stock is estimated based on the chosen estimator. Monthly spreads of each stock are calculated as the average of daily spread estimates of those stocks with at least 12 days of daily spread estimates in a given month. The mean monthly spreads of top decile stocks (largest 10% of stocks), bottom decile stocks (smallest 10% of stocks) based on market capitalisation and the mean monthly spreads across all stocks in our sample universe are reported for each type of estimator. Range-Based spread estimates are estimated from January 1973 to December 2014, and due to limited data availability closing quoted spread estimates are estimated only from January 1993 to December 2014. Data Source: CRSP.



Executive Summary

It should be noted that the numbers reported reflect full spreads (rather than half spreads). Therefore, the spread estimates reflect the average transaction costs for a round trip trade in the given universe of stocks.

The results suggest that both of our estimates provide overall similar results. The results also allow interesting conclusions to be drawn on the level, the time series variation, and the cross sectional variation of transaction costs. The average spread across all stocks had frequently reached values above 2% in the 1970s, but is situated clearly below 1% in the recent part of the sample. In such an equal-weighted average across 3,000 stocks, small stocks with high spreads obviously have a high influence. When looking at the top decile (i.e. the 300 largest stocks by market cap), the spread has taken on typical values in the area of 0.5% even during the early periods such as the 1970s. In contrast, the smallest decile stocks had historically reached spread levels exceeding 5%. We also observe spikes in the spread estimates which correspond to liquidity crisis. In particular, spikes are observed in the period from late 2008 to early 2009 – a period which saw major bank failures and a drying up of liquidity.

It is worth discussing how transaction costs behaved at points when market structure changed. In the US stock market, there are a few notable points when minimum tick sizes declined. The first change occurred in 1997 when the tick size was reduced from 1/8th to 1/16th, and the second major reduction occurred in 2001 when the tick size went from 1/16th to 1/100th (i.e. decimalisation). Smaller tick sizes allow for more competitive spreads. We can see that there is indeed a general reduction in spread levels if we compare the period prior to 1997 to the period after 2001.

Analysing Smart Beta Strategies

We apply the transaction cost estimates to several smart beta strategies to draw conclusions on their implementability. For our cost estimates, we use the closing spread estimator for the period when data is available, and the range based estimator prior to that. Our empirical analysis leads to several important conclusions in terms of replication cost estimates for smart beta strategies, which we summarise and illustrate below.

Transaction costs and implementation challenges crucially depend on the stock universe

First, we find that conclusions about transaction cost levels and strategy implementation challenges are heavily dependent on the stock universe used. While it is common to see broad brush statements about the investability hurdles of particular smart beta strategies, our results provide clear evidence that conclusions heavily depend on the universe under consideration. Our results on generic strategies show that cost metrics and investability metrics differ tremendously across universes.

A summary of results is shown in the following exhibit. We assess different universes where we select the largest 250, 500, 1,000 and 3,000 stocks to reflect different investment universes with different levels of liquidity as a starting point for implementing smart beta strategies. We then analyse portfolios drawing on random selections from these universes to assess outcomes for different weighting schemes and universe sizes chosen. To assess generic weighting schemes, we look at market cap-weighting as well as two non-cap-weighted weighting schemes, namely weighting based on firm fundamentals and equal-weighting.

Executive Summary

Implementation Costs of Generic Alternative Weighting Schemes (USA Long Term Track Records (LTTR) – Long Term – 42 Years) The time period of analysis is 31-Dec-1972 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. From the 3,000 largest stocks in the USA, universes comprising the 250, 500, 1,000 and 3,000 largest stocks are chosen and from each universe 1,000 random samples of 100 stocks are selected and weighted according the generic weighting scheme chosen. Average statistics across random portfolios are reported below. Data Source: CRSP, Compustat.

USA Long-Term 31-Dec-1972 to 31-Dec-2014	Number of Stocks in the Universe			
	250	500	1,000	3,000
Transaction Costs				
Cap-Weighted	0.04%	0.04%	0.04%	0.05%
Equal-Weighted	0.13%	0.14%	0.17%	0.38%
Fundamental-Weighted	0.11%	0.12%	0.13%	0.16%
Days to Trade (95 %ile)				
Cap-Weighted	2.06	2.39	3.79	9.99
Equal-Weighted	3.56	5.35	12.74	107.30
Fundamental-Weighted	2.48	2.93	4.80	15.13

These results underline the dependence of implementability on the universe used as a starting point. For example, for portfolios built from the top 250 stocks by market cap, we obtain Days to Trade measures of 3.56 days for equal-weighted portfolios compared to 2.06 for the cap-weighted portfolios in the same universe. Moreover, the estimate of average annualised transaction costs is 0.13% for the equal-weighted portfolios compared to 0.04% for the cap-weighted portfolios in the same universe. When looking at portfolios formed from the broad universe (the top 3,000 stocks by market cap), we get strikingly different results. The Days to Trade measure reaches more than 100 for equal-weighted portfolios compared to about 10 for cap-weighted portfolios. Estimated transaction costs are 0.38% for equal-weighted portfolios compared to 0.05% for cap-weighted portfolios. Thus an equal-weighting strategy indeed looks extremely challenging to implement for the broad universe, but implementation measures are rather well-behaved for the large cap universe. Given such differences, it makes little sense to make statements about the investability of any given strategy per se without considering the universe it is implemented for.

Practical implementation rules effectively ease liquidity and cost issues

Our analysis provides evidence of the usefulness of practical implementation rules. Our results suggest that whether or not smart beta strategies face implementation hurdles depends on the set of implementation rules that have been included in the design. We test available index strategies by comparing them to stylised portfolios that omit the implementation rules applied in practice. Our results suggest that smart beta strategies may indeed appear challenging to implement when abstracting from commonly used implementation rules, but applying these rules leads to different conclusions. For example, we report results (see the following exhibit) for a minimum volatility strategy before applying implementation rules and compare this to the same strategy after such rules have been incorporated. We show that estimated annualised transaction costs change from 0.38% to only 0.18% and investability measures such as Days to Trade go from 3.14 to 2.19 when applying practical investability rules. Perhaps more importantly, amounts traded in any stock relative to its market cap weight decline drastically from a trading multiple of 15 to

Executive Summary

Impact of Turnover and Liquidity Rules on Minimum Volatility Strategy
The time period of analysis is 31-Dec-1972 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. See Table 10 in the main part of the paper. Data Source: CRSP, Scientific Beta

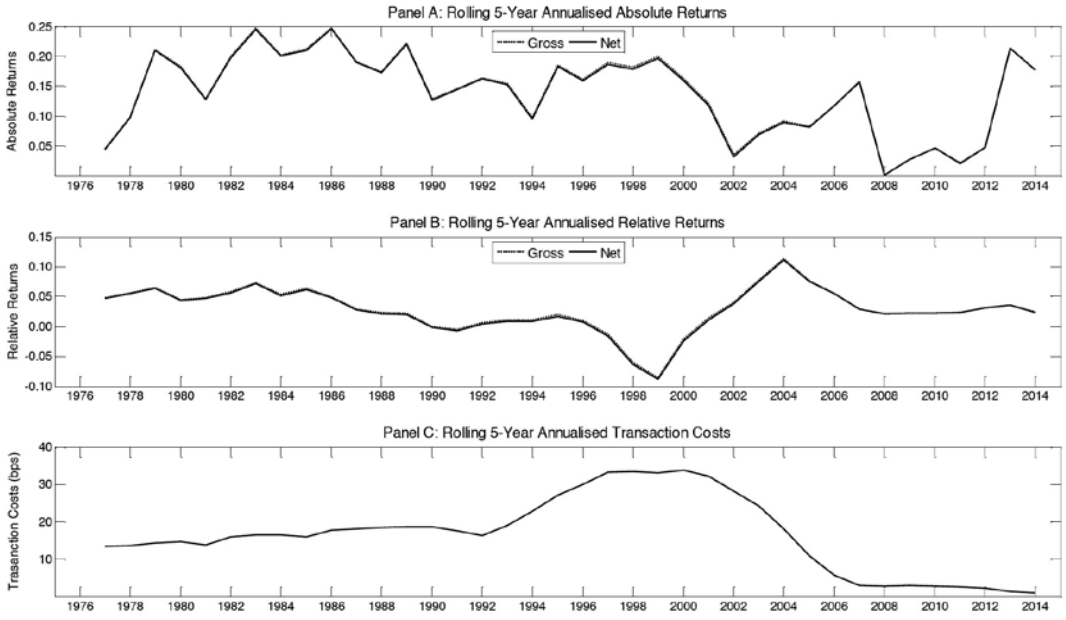
USA LTTR Long-Term 31-Dec-1972 to 31-Dec-2014	Efficient Minimum Volatility		
	Before Turnover and Liquidity Rules	After Turnover but Before Liquidity Rules	After Turnover and Liquidity Rules
One-Way Turnover	54.57%	37.96%	30.02%
Transaction Costs	0.38%	0.29%	0.18%
Days to Trade (95 %ile)	3.14	3.13	2.19
Trading Multiple (99 %ile)	15.53	9.64	1.30

a multiple of around 1. Applying common sense implementation rules thus reduces transaction costs and limits any stress on available trading volume.

Replication costs of practical smart beta strategies

We find that for the set of indices included in our analysis, which respect a set of implementation rules, smart beta performance benefits largely survive transaction costs. When looking at commonly used smart beta indices that are built on liquid universes and integrate implementation rules, the impact of transaction costs on returns is small, far from cancelling out the relative return benefits over cap-weighted indices. Transaction costs are an order of magnitude smaller than relative returns, meaning that

Rolling Window Analysis (Average across three Strategies; USA Long Term Track Records)
The exhibit presents the average annualised gross and net returns, gross and net relative returns and transaction costs of the three smart beta strategies – the SciBeta USA LTTR Efficient Minimum Volatility Index, the SciBeta USA LTTR Maximum Deconcentration Index and the SciBeta USA LTTR Multi-Beta Multi-Strategy (4-Factor) EW Index using a rolling 5-year window with 1-year step size. Panel A presents the gross and net absolute returns; Panel B presents the gross and net relative returns; Panel C presents the transaction costs. The average returns/costs of the three smart beta indices each year are plotted. The time period of analysis is 31-Dec-1972 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. The transaction costs estimates use the spread estimates according to the year of the rebalancing – Range-Based spread until 1993 and Closing Quoted Spread from 1993 onwards. The reported transaction cost estimates are the difference between the annualised gross and net returns. Net returns are obtained after accounting for transaction costs at each quarterly rebalancing by multiplying the change in weight of each stock between the final weight before rebalancing and the optimal weights after rebalancing, including stock deletions and additions.



Executive Summary

net relative returns do not differ materially from gross relative returns. For the three strategies we consider, namely a Minimum Volatility, Maximum Deconcentration and a Multi-Factor index, we find that average annualised transaction costs over the 42-year period are between 0.13% to 0.18%, while gross returns relative to the cap-weighted index range from 2.38% to 3.93%. The exhibit on the previous page shows five-year rolling window returns, both net and gross returns. For brevity, the graphs show the average return across the three strategies analysed. It is rather clear from inspecting the lines for net and gross returns that transaction costs hardly alter the returns of these strategies. However, it should be noted that such a conclusion cannot hold for smart beta strategies in general, as emphasised in our first two findings. For example, with a less liquid universe or less stringent implementation rules, the same strategies may be burdened by much higher transaction cost levels and implementability issues.

Managing switching costs into smart beta strategies

Another aspect which is important to analyse is the potential cost of switching into smart beta strategies, when investors replace a currently invested portfolio with a new strategy. As a reasonable starting point from which the switch occurs, one can assume a cap-weighted portfolio based on the underlying index universe. It should be noted that investors can manage the cost of switching from cap-weighted indices to smart beta strategies in a straightforward way by stretching out the transition from a cap-weighted portfolio to a smart beta strategy. In the following exhibit, we address both the transaction costs that occur through rebalancing and those that occur when initially switching from a cap-weighted index to the smart beta strategy. In order to estimate switching costs for a 10-year investment period, we apply trading cost estimates to the trades needed to switch from the cap-weighted index to the smart beta index and compute the corresponding annualised costs assuming that the switch is done for a subsequent investment period of 10 years.

It can be seen that the stretching the transition over a period improves the Days to Trade but the returns remain almost the same. The tracking error between the stretched and non-stretched portfolios also remains quite low although they increase in the stretch period. The cost of transition is very small compared to the cost of rebalancing and the total cost is still low compared to the gross returns even after accounting for the transition costs.

Executive Summary

Comparison of stretched and non-stretched transition from Cap-Weighted portfolio to Smart Beta Portfolio (Long Term - 42 years) The time period of analysis is 31-Dec-1972 to 31-Dec-2014. The strategies considered for this analysis are the SciBeta USA LTTR Efficient Minimum Volatility Index, the SciBeta USA LTTR Maximum Deconcentration Index and the SciBeta USA LTTR Multi-Beta Multi-Strategy (4-Factor) EW Index. All statistics reported in Panel A are quarterly estimates and are averaged across all quarters. Results of three types of scenarios are estimated and presented – i) The switch from Cap-Weighted portfolio to Smart Beta portfolio happens completely on the day of rebalancing (1-day Transition); ii) The switch from Cap-Weighted portfolio to Smart Beta portfolio happens equally distributed across 10-days (10-day Transition i.e. assuming only one-tenth of the portfolio switches every day for 10 days); iii) The switch from Cap-Weighted portfolio to Smart Beta portfolio happens equally distributed across 20 days (20-day Transition i.e. assuming only one-twentieth of the portfolio switches every day for 20 days). Days to Trade (DTT) is reported as a time-series average of the cross-sectional 95th percentile of DTT at each quarterly rebalancing. Tracking Error of stretched transition (both 10-days and 20-days) over non-stretched transition is computed quarterly and average is reported. Difference in Gross Returns is computed quarterly between stretched (both 10-days and 20-days) transition and non-stretched transition. All statistics reported in Panel B are annualised. It compares costs of all three smart beta strategies. Assuming 10 year investment period, the Annualised Cost of Transition from Cap-Weighted Index is computed as one-tenth of the immediate transition (a semi-absolute difference between weights of smart beta strategies and Cap-Weighted index multiplied by the average weighted spread and averaged across all quarters). Annualised Cost of Rebalancing is the average difference between annualised gross and net returns. Total Annualised Cost is sum of transition and rebalancing costs.

USA LTTR Long-Term 31-Dec-1972 to 31-Dec-2014	Transition	Efficient Minimum Volatility	Maximum Deconcentration	Multi-Beta Multi-Strategy 4-Factor EW
Panel A: Transition from Cap-Weighted Index (statistics for transition quarter)				
Days to Trade (95%ile)	Non-stretched	1.72	1.98	2.64
	Stretched 10-days	0.17	0.20	0.26
	Stretched 20-days	0.09	0.10	0.13
Tracking Error	Non-stretched	-	-	-
	Stretched 10-days	0.08%	0.08%	0.09%
	Stretched 20-days	0.12%	0.11%	0.12%
Difference in Gross Returns by Stretching	Non-stretched	-	-	-
	Stretched 10-days	0.00%	0.01%	0.00%
	Stretched 20-days	-0.01%	0.00%	-0.01%
Panel B: Cost Comparison				
Annualised Cost of Transition from Cap-Weighted (assuming 10-year investment period)		0.02%	0.02%	0.03%
Annualised Cost of Rebalancing		0.18%	0.13%	0.17%
Total Annualised Cost		0.20%	0.15%	0.20%

Conclusions

The results in this paper provide an important contribution to the analysis of smart beta strategies from a practical perspective. Indeed, the state of affairs in the evaluation of smart beta strategy performance is far from satisfying. On the one hand, strategy providers do not commonly report the transaction cost estimates of their strategies, and performance evaluation often relies on simulated gross returns. On the other hand, the discussion of cost issues more often than not remains at the level of blanket criticism aimed at certain strategies, without considering the universe or the implementation rules that are used. Our results provide an explicit estimate of costs applied to a range of strategies and show the impact of using different implementation rules or stock universes. Importantly, given the transparent methodology and benign data needs, our replication cost analysis is straightforward and can be easily applied to other strategies.

Introduction

Introduction

The objective of this paper is to provide estimates of costs that arise when replicating smart beta index strategies. Such cost estimates allow us to report the net returns of such strategies. In fact, while the importance of accounting for transaction costs in the evaluation of such strategies is not doubted, relatively little is known about the magnitude of these costs. Consequently, relatively little is known about the impact of implementation aspects, such as the selection of a liquid universe or turnover controls, on transaction cost levels.

Smart beta strategies are commonly analysed on the basis of backtested performance. Typical backtest results do not consider real life transaction costs. Index providers sometimes state that adjusting the performance of their indices for transaction costs is not something that they can do easily and they prefer to leave it to market participants to figure out what the shortfall might be. When providers do address this issue, they often rely on arbitrary assumptions about transaction cost levels (such as fixed costs irrespective of the time period or the type of stock that is traded). In both cases, there is not even an attempt to provide explicit estimates of costs.

In similar vein, research results introducing particular smart beta strategies rarely contain estimates of transaction costs. It is common to present the outperformance of smart beta strategies over cap-weighted indices without subtracting the transaction costs due to portfolio rebalancing (Asness, Moskowitz and Pedersen, 2013; Blitz and van Vliet, 2007;, Chan, Karceski and Lakonishok, 1999; Choueifaty and Coignard, 2008; Maillard, Roncalli and Teïletche, 2010). In some cases, there is not even discussion as to the possible impact of transaction costs on alternative portfolio performance. In other cases, authors explicitly recognise that they ignore transaction costs or use qualitative arguments to claim that transaction costs should not alter results. In some cases, portfolio turnover is reported as a rough indicator of implementation hurdles. Other papers do not use explicit cost estimates but provide naive proxies. Some papers use the turnover and an assumption of constant costs across all trades to calculate indifference levels of fixed costs that would cancel the outperformance of a strategy (Arnott, Hsu and Moore, 2005; Amenc et al., 2011). Such attempts at providing some information on the cost dimension are an improvement, compared to papers that do not take transaction costs into account at all. However, these papers are still short of an explicit estimate of actual costs. In particular, assuming that costs are constant across all trades is highly unrealistic as it is obvious that cost levels will vary across stocks and over time. This state of affairs may not be entirely satisfying as, in the absence of considering transaction costs, the performance of smart beta indices will be naturally overstated.

On the other hand, some market participants make bold claims that even some popular strategies cannot be implemented without a substantial performance drag from transaction costs. A common claim is that equal-weighted strategies cannot be implemented without incurring excessive transaction costs. For example, West (2007) argues that "*turnover is relatively high and expensive for an equal-weighted portfolio*", because "*transactions would occur in relatively small, less liquid stocks*" and "*given these drawbacks, it is clear that a simple equal*

Introduction

weighting approach is not an attractive investment strategy for most investors." Unfortunately, these statements are not accompanied by actual measures of transaction costs. However, in the absence of transaction cost estimates for such strategies, such claims cannot be evaluated.

As previously stated, the objective of this paper is to assess transaction costs of smart beta strategies so as to contrast the gross returns of such strategies shown in backtests with estimates of net returns that are actually available to investors when considering transaction costs. A first important objective of our research is to document transparent methods of estimating transaction costs based on widely available data that can be used to improve the information available to potential smart beta investors. A second objective is to apply such estimates to a set of different smart beta strategies to provide insights into the resulting replication cost levels.

Research results containing estimates of transaction costs for some smart beta strategies are available. However, results are often limited to particular strategies and/or based on cost estimation methodologies which may be limited or proprietary. Two common cost approximations are obtained through plug-in methodologies and through simple extrapolation. In plug-in approaches, one simply looks up pre-recorded cost information without transparency on its determinants. In extrapolation approaches, one draws on reported parameters from studies that have run regressions linking observed high frequency transaction costs to stock characteristics in a given sample.[1]

For example, Auer and Schuhmacher (2015) plug in historical transaction cost records for stocks and apply them to one particular strategy. Chow et al. (2011) consider a range of alternative weighting schemes but use cost estimates which they qualify as "rough references only" based on extrapolation. De Groot, Huij and Zhou (2011) assess a particular strategy by plugging in values from a proprietary transaction cost data source, and compare this to the extrapolation approach. Frazzini, Israel and Moskowitz (2015) analyse factor based equity strategies drawing on a proprietary transaction cost database.

In contrast to such approaches, Novy-Marx and Velikov (2016) use explicit cost estimates drawing on empirical microstructure models. Our paper is similar in that it also uses explicit cost estimates based on such models. However, our paper differs from theirs in that we test popular smart beta strategies while they test academic factor portfolios that are quite different from popular smart beta strategies. Moreover, their results are based on a methodology with limited empirical performance (Hasbrouck, 2009). We draw on recent advances in microstructure research that allows us to extract reliable estimates of transaction costs from low frequency (daily) data. Our estimation methods have been shown to outperform, for example, the Hasbrouck (2009) estimate used by Novy-Marx and Velikov (2016) in terms of reliably reflecting transaction cost measures that are obtained from intraday data.

Moreover, our estimation methods draw on easily available data and are not computationally intensive, making our results easily replicable. We apply our cost estimates to a range of smart beta

1 - A popular set of parameters comes from Keim and Madhavan (1997) who analyse a sample of stocks for a period of about two years.

Introduction

strategies which are widely use in practice and our transaction cost measurement approach can easily be applied to testing additional strategies.

We follow two types of spread estimation methods based on daily data – one based on Corwin and Schultz (2012) who use daily range measures such as high and low prices to estimate daily spreads (hereinafter referred to as the *range-based* spread estimator), and the other based on Chung and Zhang (2014) who use daily closing quoted bid and ask prices to estimate daily spreads (hereinafter referred to as the *closing spread* estimator).

While the performance of these estimators in capturing information on transaction costs that would be available from high frequency data analysis has been widely documented and confirmed in the literature, we first compare our estimates to a set of benchmark values. Relative to transaction cost estimates from intraday data, we observe that both the range-based spread measure and the closing spread provide decent proxies in terms of level and in terms of time series variation.

When applying these cost estimates to different smart beta strategies, we find that transaction costs differ widely depending on how a strategy is implemented. Strategies applied to easily tradable stocks and subjected to practical implementation rules have net returns that are hardly distinguishable from gross returns, while transaction costs have a larger impact for strategies that draw on less liquid universes or those that omit implementation rules. Our results thus show the importance of considering strategy implementation.

The remainder of this paper proceeds as follows. Section 1 introduces our methodology for computing transaction costs, drawing on recent advances in low frequency microstructure research. Section 2 presents our transaction cost estimates and compares them with alternative estimates including those obtained from high frequency trade and quote data. Section 3 applies these cost estimates to smart beta strategies to estimate net returns. A final section concludes.

1. Methodology

1. Methodology

This section introduces our methodology for estimating transaction costs. We start by providing a review of possible methodologies of interest and explain our choices.

1.1 Trading Cost Definitions

Trading costs lower the price for sellers of securities and increase the price for buyers, relative to a market where trading would be frictionless. Total trading costs can be decomposed into three components: the spread, the price impact and commissions (see e.g. Hasbrouck (2007) or Connor, Goldberg and Korajczyk (2010, Chapter 12) for an introduction to trading costs).

The spread reflects the cost of a round trip trade (buy and sell). The basic cost definition is the per cent quoted spread which reflects the percentage cost of buying at the ask quote and selling at the bid quote. If trades occur at prices different from the best bid and offer (BBO), the per cent effective spread is a more useful measure. This measure is based on the (absolute) deviation of the transaction price from the frictionless price (the bid-ask midpoint). The deviation is then multiplied by two so as to reflect round trip costs. If trades only occurred at the BBO, this measure would yield the same result as the percentage quoted spread, but in the presence of trades at prices different from the BBO, the effective spread indicates the actual round trip trading costs. In particular, large orders will create a price impact which is captured by the effective spread. For example, a large buy order may lead to an increase of the price beyond the ask quote displayed when the trade was initiated. Such price impact is captured by the effective spread. Huang and Stoll (1996) provide a method for decomposing the effective spread into a price impact and so-called realised spread, with the latter reflecting the cost of order processing and inventory risk. Spread measures in principle can be derived from observing the intraday trades and quotes, and from calculating volume weighted averages across all trades occurring for a security within a trading day.

While the effective spread includes a price impact component, this is the price impact given the observed trading volume and trading behaviour. Obviously, investors who trade aggressively and act as liquidity demanders may incur a higher price impact than the average price impact reflected in the empirical effective spread estimate. Similarly, investors who trade smartly may incur lower transaction costs than the average estimate reflected in trade and quote data. It would be interesting to assess such variations in trading costs by coming up with actual models for price impact, also called-cost curves. However, there is much debate in the literature on the correct specifications of such curves and the resulting magnitude of transaction costs (see Frazzini, Israel and Moskowitz (2015) for a recent example within this discussion). In particular, price impact is difficult to estimate when extrapolating outside observed volume ranges, and there is ample discussion but no consensus on the appropriate functional form of price impact functions. For these reasons, it is beyond the scope of this paper to discuss precise cost estimation depending on different order strategies. This may however be an interesting avenue for future research. It should be noted, however, that our measures do capture the price impact component related to the observed trading volume and order strategies, in the sense that they aim to be good estimates of the effective spread.

1. Methodology

Another source of transaction costs are commissions that brokers may charge for routing orders to market places. However, such commissions are not visible in quotes and transactions, and may be difficult to account for since commission schedules may vary widely depending on contractual arrangements and bargaining power of investors.[2]

Our main focus in this study is to capture the effective spread component because it can be reliably estimated from market data. Even when it comes to the spread component, studies smart beta transaction costs are rare because accurate estimation in principle requires intraday high frequency data. As mentioned above, we need to observe trades and quotes within the trading day to come up with cost measures. However, such data is both hard to use and hard to get. It is hard to use because high frequency data is big and messy. The increasing frequency of trading has led to a huge amount of tick by tick price data that requires massive computational power for analysis. Fong, Holden and Trzcinka (2014) who analyse several billion data points argue that high frequency equity data likely grows at a rate of more than 30% per year which outpaces the growth of computing power. Moreover, tick data requires matching procedures for prices and quotes (Lee and Ready, 1991), as well as intense data cleaning so that the quality of databases and the cleaning procedures becomes a prime concern. High frequency data is hard to get because it is expensive, and it covers only short time periods. It is common for researchers to analyse periods of less than a decade, and sometimes only a few years, due to data availability limitations.

Recent research has shown that there are effective ways of estimating transaction cost variables that are only observable at high frequency, based on lower frequency (daily) data. The advantage of such approaches is that results can be generated for longer periods and different markets, with relative computational ease and limited data needs.

1.2 Low Frequency Measures

There is a fairly broad set of competing methods that allow estimates of effective spreads to be extracted from daily data. Such methods produce a spread proxy for a given stock in a given month. The common basis of most such measures is that transaction costs produce a distortion of the true price process that would prevail without any transaction costs. Thus observable prices are influenced by transaction costs. Low frequency transaction cost measures attempt to extrapolate the transaction cost component from observable prices. To do this, the different measures use information from various properties of the observed price path, such as its serial dependence, occurrence of zero returns, observed price clustering or the observed price range.

The first estimator was proposed by Roll (1984) and is based on serial dependence in stock returns. Roll postulates that the closing price will reflect the true price inflated by a half spread if it is buyer-initiated, and deflated by the half spread if it is seller-initiated. Roll derives that, for a price process with zero serial correlation, the distortion brought in by the bid-ask spread will introduce negative serial correlation, as the price will tend to bounce back from a bid price when the next trade is at the ask price, and vice versa. Under rather stringent assumptions, notably assuming that trade direction is not serially correlated, the bid-ask spread is just

2 - We also note that, other than commissions, our analysis will not include cost categories such as financial transaction taxes, exchange taxes, or a stamp duty which may exist in certain markets. Moreover, taxes that may be applicable at the investor level (e.g. taxes on realised capital gains) that may also differ depending on rebalancing schedules and implementation approach are not considered. In fact, such taxes may vary depending on the status of the investor which would make it challenging to include them in the analysis.

1. Methodology

a function of the serial covariance. More specifically, the half spread is expressed as $\hat{S}_{Roll} = \sqrt{-Cov(\Delta P_t, \Delta P_{t-1})}$, where ΔP_t refers to the price change observed on day *t*. However, using the Roll estimator is problematic if empirical serial covariance is positive or if there are days where stocks do not trade. Hasbrouck (2009) develops and tests a Bayesian approach to estimating spreads in the Roll model.

Lesmond, Odgen and Trzcinka (1999) develop transaction cost estimates based on observing returns that are zero. Daily returns that are zero actually provide information on the absence of informed trading due to high transaction costs. When true returns are within the range given by the transaction costs for buying or selling the stock, we will observe a zero return. Based on this insight, Fong, Holden and Trzcinka (2014) propose a simple measure based on the percentage of days in a month with zero returns.[3] Their estimate of the half spread is defined as $\hat{S}_{FHT} = \sigma N^{-1}\left(\frac{1+z}{2}\right)$ where *z* denotes the number of days with zero returns as a percentage of the total number of days in a month (including non trading days), σ denotes the daily return volatility of the stock and N^{-1} is the inverse of the normal cumulative distribution function. We can see that the estimated spread is increasing in volatility and in the frequency of zero returns. The reason for this is intuitively plausible: Given the same level of return volatility, a stock with higher transaction costs will display zero returns more frequently than a stock with low transaction costs. Likewise, given the same percentage of zero returns, a stock with higher volatility must have higher transaction costs.

Other measures are based on clustering of observed prices. Holden (2009) and Goyenko, Holden and Trzcinka (2009) develop the Effective Tick measure. They assume that the spread size determines how observed prices are clustered at levels that ease negotiating between traders. Thus the frequency of closing prices observed for particular clusters allow backing out the spread size. Holden (2009) combines information from such price clustering with information on serial correlation to obtain yet another spread proxy. However, these measures are "very numerically-intensive" and it becomes "infeasible" to apply them to a large sample of stocks (Fong, Holden and Trzcinka, 2014).

It has been shown that such low frequency measures reliably proxy for the effective spread as measured from intraday trade and quote data. Cost measures that perform well empirically need to have high time series correlation, high cross sectional correlation, and low root mean squared error with the effective spread obtained from high frequency data. Goyenko, Holden and Trzcinka (2009) conduct an extensive comparison of different spread proxies with spreads derived from intraday trade and quote data for US listed stocks. They conclude that the low frequency spread proxies capture both the cross sectional variability and the magnitude observed in spreads obtained from high frequency data. They conclude: "*The evidence is overwhelming that [...] low-frequency measures capture high-frequency measures of transaction costs.*" When comparing across the different measures, they conclude that spread proxies based on price clustering and zero returns capture high frequency information more reliably than the other measures.

While there is evidence that the different measures are good proxies of transaction costs, recent advances in the empirical microstructure literature have introduced

3 - Lesmond, Odgen and Trzcinka (1999) derive a maximum likelihood estimator of transaction costs in an explicit model of stock returns and transaction costs based on this intuition, referred to as the LOT measure. Goyenko, Holden and Trzcinka (2009) propose several different estimation methods for the LOT model.

1. Methodology

new ways of estimating transaction costs from low frequency data which dominate the earlier ones in terms of ability to proxy for high frequency estimates of transaction costs. We are particularly interested in using these newer methodologies that have been shown to work extremely well empirically. It is worth noting that the relevant benchmark for comparison used when establishing the quality of low frequency spread estimates is the *effective spread* from high frequency data, which includes both a realised spread component and a price impact component.

In particular, Corwin and Schultz (2012) develop a method which measures bid-ask spreads based on the daily price range (high and low prices within the trading day) and Chung and Zhang (2014) use the bid-ask spread quoted at the close of the day as an estimate. Corwin and Schultz (2012) and Chung and Zhang (2014) provide evidence that their new spread measure provides more reliable estimates of effective bid-ask spreads compared to other low frequency methods. In particular, they provide evidence that the cross sectional and time series correlation of their respective measure with the transaction cost information obtained from intraday trade and quote data are consistently high. Fong, Holden and Trzcinka (2014) conduct an extensive empirical comparison of different low frequency methods and their capacity to capture the information about transaction costs from high frequency trade and quote data from more than 40 exchanges around the world. They conclude that the closing quoted spread and the range-based spread dominate other low frequency measures in terms of capturing information in high frequency data. Importantly, they provide evidence that such low frequency measures do not only capture the cross sectional and time series variation of transaction costs, but also their magnitude. Abdi and Ranaldo (2017) conduct a similar comparison where the range-based estimator of Corwin-Schultz and the closing spread estimator also fare better than the other estimators discussed above.[4] It should also be noted that these two measures can be computed at daily frequency while earlier measures take daily data as inputs but can only be computed at monthly frequency. Based on such findings, several recent papers use these methods to derive spread measures when accounting for trading costs of equity factors (McLean and Pontiff, 2016) or of trading strategies such as merger arbitrage (Giglio and Shue, 2014). Beyond stock markets, several authors find that the same low frequency measures are able to successfully capture information on transaction costs in other markets such as the corporate bond, (Schestag, Schuster, and Uhrig-Homburg, 2016), currency (Karnaukh, Ranaldo and Söderlind, 2015) and commodity markets (Marshall, Nguyen and Visaltanachoti, 2012).

In the following two subsections, we explain the two methodologies which we employ in our tests in greater detail.

1.2.1 The closing quoted spread

Definition

Without access to intraday trade and quote data, one can still obtain an estimate of the prevailing spread from information that is available at the close of the trading day. Chung and Zhang (2014) introduce this measure into the literature and analyse the closing spread measure for US equities based on CRSP data. The closing percentage half spread $\hat{S}_C$ is calculated as follows:

$$\hat{S}_C = \frac{(a_c - b_c)}{(a_c + b_c)}$$

where a_c is the closing ask price and b_c is the quoted bid price at close. Thus $\hat{S}_C$ is

4 - They also introduce a new estimator based on both the daily range and the closing price. We compared their estimator with the range-based estimator we use and concluded that there are no tangible differences in terms of the results in our dataset. This is consistent with their own findings where they show differences between the two estimators essentially when it comes to low liquidity stocks with high spreads which are not included in the investible universes we focus on in this paper.

1. Methodology

the half-spread and $\widehat{2S_C}$ is the full spread, expressed in percentage terms relative to the bid-ask midpoint $(a_c+b_c)/2$.

Chung and Zhang (2014) provide a detailed discussion of the corresponding CRSP data items. The closing ask and bid quote are available from the CRSP database consistently for both NYSE/AMEX and NASDAQ listed stocks only from 1993 onwards. For NASDAQ stocks the closing bid and ask are the best quotations (highest bid and lowest ask), while for NYSE/AMEX stocks the closing quotes are the last quotes before the market close.

Fong, Holden and Trzcinka (2014) use a similar measure for international equities based on Thomson Reuters Datastream data. They find that the required data items, closing bid and closing ask quotes, are widely available across a range of global stock exchanges in Datastream with good coverage from the late 1990s onwards. The measure is thus broadly applicable to analyses of equity investment strategies across different countries.

Adjustments

We follow Chung and Zhang (2014) and make the following adjustments. First, CRSP sets the closing bid or ask to zero if the last quote is not representative of actual trading activity. In these cases, we treat the data as missing. We also exclude any closing spread estimate that exceeds 50% to avoid inclusion of erroneous data points.

As the data limitations mean that we have this spread measure only for the recent period, the range-based estimator is a much more general measure and will be especially useful for longer samples extending back before 1993. We now turn to a discussion of the range-based spread measure.

1.2.2 The range-based spread estimate

This subsection introduces the high-low spread estimate and provides illustrations on the principles underlying its derivation.

Definition

The ratio of the daily high to the daily low price allows us to obtain information on spreads for a simple reason. This ratio depends on the volatility of the true stock price, as well as the bid-ask spread. If one can disentangle these two components, it is possible to compute the spread component.

The high price will occur for a buyer-initiated trade and will correspond to the true price plus half a spread. The low price will occur for a seller initiated-trade and will reflect the true price minus half a spread. It is then straightforward to see that the observed high/low range depends on the range of the true price and the bid-ask spread. In particular, with h_t and l_t denoting the observed high and low prices, H_t and L_t denoting their true price counterparts and $2S_t$ the true spread we have

$$\frac{h_t}{l_t} = \left(\frac{H_t(1+S_t)}{L_t(1-S_t)}\right).$$

Taking natural logarithms on both sides, we see that the observed price range has two components – the range of the true price, and a microstructure component due to the spread:

$$ln\left(\frac{h_t}{l_t}\right) = ln\left(\frac{H_t}{L_t}\right) + ln\left(\frac{1+S_t}{1-S_t}\right).$$

The key idea underlying the range-based spread estimator is to compare the (observed) high-low ratios for two single-day periods to the high-low ratio over the two-day period. The high-low ratios over two one-day periods reflect two spreads, and two one-day volatilities. The high-low ratio over a single two-day period reflects one spread and the volatility over two

1. Methodology

days. Furthermore, the two-day volatility is proportional to one-day volatility, and this relation allows the spread component to be isolated.

The final spread proxy is derived by solving a system of equations, using results from Garman and Klass (1980) and Parkinson (1980) on the relationship between high-low ranges and volatility which can be expressed in closed form when stock prices follow Geometric Brownian motion.

Key components in the final proxy for the effective spread are the sum of two daily high-low ratios β and the high-low ratio over a single two-day period γ (both in terms of expected squared log price ratios):

$$\beta = E\left[\left(ln\left(\frac{h_t}{l_t}\right)\right)^2 + \left(ln\left(\frac{h_{t+1}}{l_{t+1}}\right)\right)^2\right],$$

$$\gamma = E\left[\left(ln\left(\frac{h_{t,t+1}}{l_{t,t+1}}\right)\right)^2\right].$$

The range-based estimator $\hat{S}_R$ for the half-spread is then given as:

$$\hat{S}_R = \frac{(e^{\alpha} - 1)}{1 + e^{\alpha}}$$

where

$$\alpha = \frac{\sqrt{2\beta} - \sqrt{\beta}}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}})$$

Adjustments

Using the method in practice requires several adjustments which are discussed in Corwin and Schultz (2012). First, the range needs to be adjusted for overnight price changes in case the closing price on day *t* lies outside the high-low range observed on day *t+1*. If the day *t* closing price is higher than the high price of day *t+1*, the latter is set to be equal to the day *t* closing price for the spread computation. In a similar fashion, the day *t+1* low will be set to the previous day's closing price if that closing price is lower. Second, the empirical daily spread measure can be negative. To avoid using negative spread estimates, the monthly spread measure is obtained as the average across all daily measures, setting any daily estimates that are negative to zero before taking the average.[5]

A key advantage of this cost measure is that it is valid under fairly general conditions and can be generated easily based on widely available data and without a need for computationally intensive procedures. Corwin and Schultz (2012, p. 721) argue that this measure also captures price impact, because large or repeated orders are often executed at the daily high (in the case of buyer-initiated orders) or low (in the case of seller-initiated orders) price. They show that the measure outperforms a wide range of alternative low frequency measures in terms of capturing information with respect to transaction cost measures extracted from high frequency data.

Illustration on the range-based spread measure

We provide a brief illustration to build some intuition on how the range-based spread estimate works in principle. Indeed, the power of this approach is that we are able to extract estimates of transaction costs without actually observing transaction costs explicitly, as we could have done for example by having access to intraday trade and quote data.

The key explanation of why we can extract estimates of transaction costs from the daily price range is that transaction costs affect the price process and the range allows us to obtain information on this process. The relation between the daily price range and volatility is well understood. Without any microstructure effects (i.e. with a zero bid-ask spread) the range

5 - Corwin and Schultz (2012) provide simulation results suggesting that replacing daily negative spread estimates by zero before averaging these measures to obtain a monthly spread estimate is preferable to averaging all daily measures (including negative ones) and then setting negative monthly spread measures to zero. Abdi and Ranaldo (2016) discuss reasons for replacing daily negative values by zero before averaging rather than applying this replacement to the monthly measures (see page 12). In particular, the range-based spread measure assumes constant volatility and constant spreads. For the two day time window used to derive the daily measures, this assumption is less likely to be problematic. Therefore, adjusting anomalous values resulting from the two-day measurement period leads to better spread estimates.

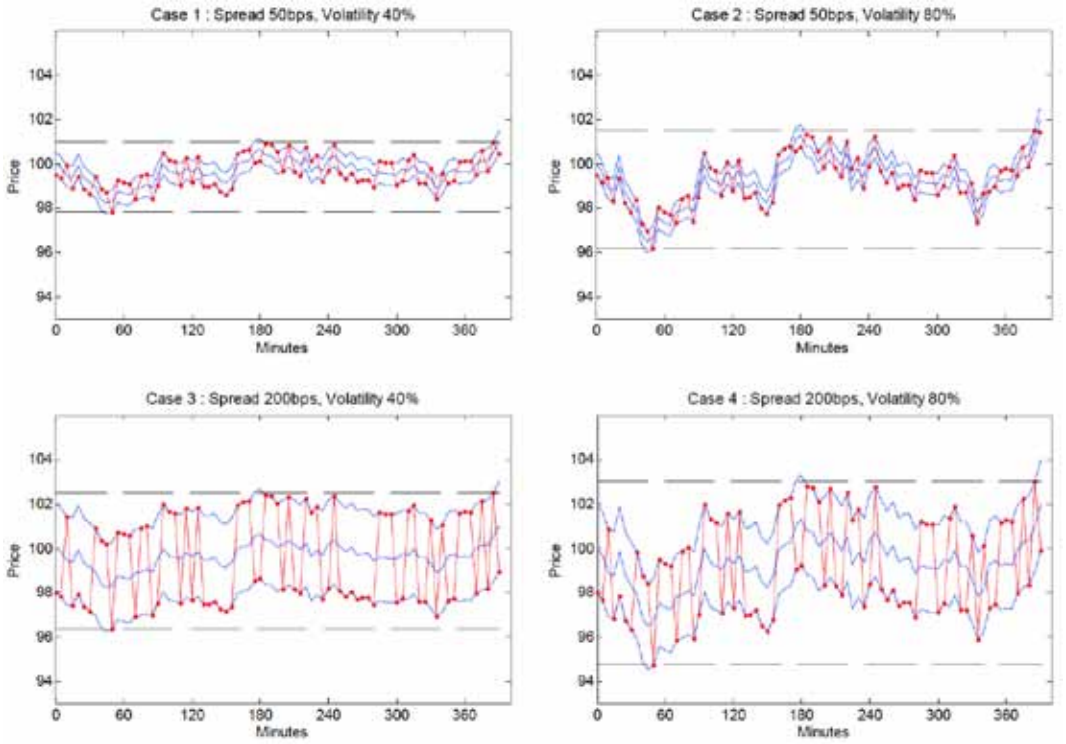
1. Methodology

would solely depend on volatility of the true price. Range-based volatility estimators have been proposed by Garman and Klass (1980) and Parkinson (1980). The bid-ask spread distorts the price process, and the range of the observed price path allows us to extrapolate the magnitude of the distortion.

To illustrate the dependency of the range on the effective spread, we generate random scenarios. We simulate the price process at every minute as in Corwin and Schultz (2012) as $P_m = P_{m-1}e^{\sigma x}$ where x is drawn randomly from a unit normal distribution and σ denotes per minute standard deviation. Moreover, a trading day has 390 minutes and we assume trades occur every five minutes at the true price deflated (seller-initiated trade) or inflated (buyer-initiated trade) by the half-spread, where we draw trade direction randomly with 50% probability.

To illustrate the effect that volatility and spread have on the daily price range, we look at four cases. We set annualised volatility to 40% and 80%, which corresponds to daily volatility of approximately 2.5% and 5%. We set the spread to 0.5% and 2%. In Exhibit 1, we plot the true price, bid price, ask price, and the observed price, and we indicate the high and low. Exhibit 1 just displays results from one random draw. It provides a clear illustration of how the range increases with increasing volatility, when comparing the plot for Case 1 and Case 2 for example. The effect of an increasing spread is shown when comparing Case 1 and Case 3 for example.

Exhibit 1: Illustration: Impact of Volatility and Spread on Daily Range
The figure shows illustratively how volatility of the unobserved true price and the spread influence the intraday range. The plots show the price paths at five minute intervals over a trading day (390 minutes) resulting from a single scenario on randomly drawn price shocks and random trade direction at each minute. Four cases are shown reflecting different parameters for volatility and for the spread. The blue lines represent (from top to bottom) the ask price, bid-ask midpoint and bid price. The red line represents the transaction prices, which take place either at the bid (seller-initiated trade) or at the ask (buyer-initiated trade).The graphs show that the range increases with the spread and with volatility of the true price.



1. Methodology

At a given level of volatility, an increase in the spread increases the daily range, roughly by the amount of the spread (Case 3 versus Case 1, and Case 4 versus Case 2).

Since volatility is itself unknown, in empirical work we cannot directly compare the observed range to the range that would result from volatility alone without microstructure effects. However, the range-based spread estimator uses the knowledge we have on the relationship between one-day volatility and two-day volatility, and the impact of volatility on the range, so as to derive the spread estimate from the high and low prices over two consecutive days.

1. Methodology

2. Descriptive Statistics on Transaction Costs

2. Descriptive Statistics on Transaction Costs

In this section, we provide a description of the transaction cost estimates we obtain for our stock universe. These are the estimates that will be applied to the returns of smart beta strategies in Section 3 to obtain net returns. In order to assess the reliability of the estimates, we compare them to a range of benchmark values that have been published in the literature or that are available in regulatory reporting or published datasets. We also asses the relationship of average transaction costs over time with exogenous variables that proxy for market conditions. Finally, we will conclude our analysis with our choice of a suitable estimation procedure for the subsequent analysis.

2.1 Transaction Cost Estimates

We first provide an overview of the magnitude of transaction costs we obtain for long-term US data. To construct the underlying stock universe, we proceeded as follows.

Our data sample is taken from the CRSP database and consists of all US ordinary common shares (i.e. shares with share code 10 or 11) listed on the NYSE, Amex and NASDAQ stock exchanges. We further filter our sample to include only the 3,000 largest stocks. The 3,000 largest stocks are selected on every rebalancing day based on the stocks' total market capitalisation as of the previous day. All the smart beta indices discussed in this document are rebalanced quarterly on the third Friday of March, June, September and December every year. The data sample is extended as far back as possible where at least 3,000 stocks are available in the three stock exchanges combined. Thus, our sample period is from December 1972 to December 2014 (42 years).

For our US universe of 3,000 stocks, we compute both the range-based spread estimates and the closing spread. Estimates are monthly estimates obtained by averaging the daily estimates for the respective month. Only stocks with at least 12 days of spread estimates in a given month are included in the analysis consistent with the approach followed by Corwin and Schultz (2012). The starting date for the two spread estimates is different. The range-based spread estimate requires daily high and low data, which in principle are available even for periods extending far back in history. The closing spread requires available data on the last bid and ask quote of the trading day. The closing quoted bid and ask prices of stocks for all the exchanges are only consistently available from 1993 in the CRSP database.[6] Hence the closing quoted spread estimates that rely on bid and ask information are computed only from 1993.

Exhibit 2 shows results for the average spread, 5th and 95th percentile across all stocks, as well as the across the largest and smallest stocks in our universe. Large and small stocks are taken as the top and bottom deciles every year by market capitalisation (as of the last trading day of the previous year). The 3,000 stocks available in every quarter of a given year are aggregated for the decile selection. Thus the number of unique stocks may be greater than 3,000 in a given year. Monthly average spread estimates are then calculated for these deciles.

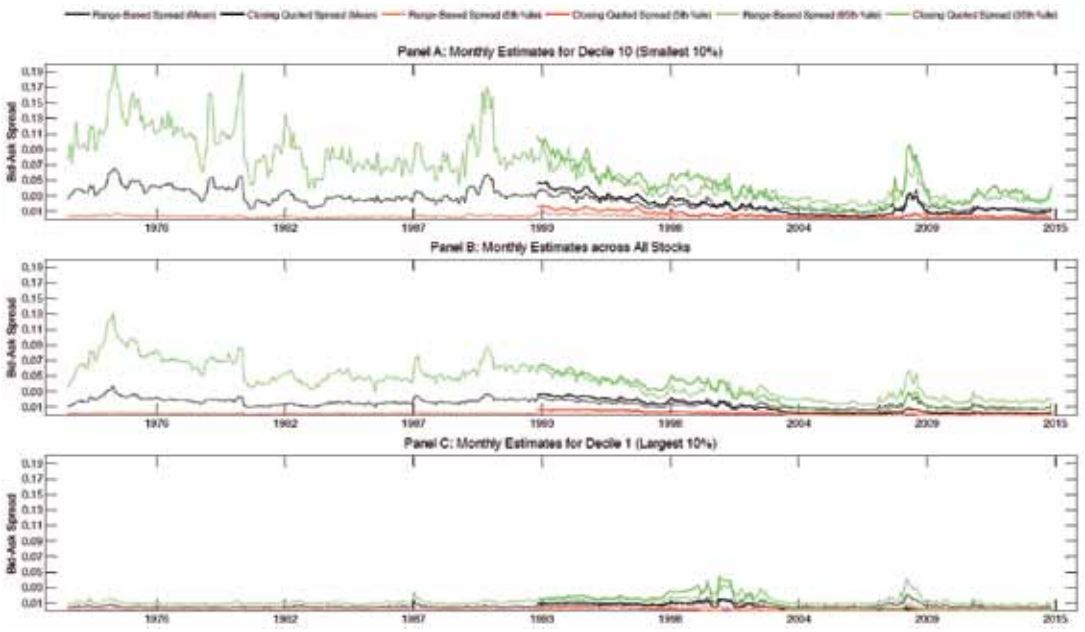
It should be noted that the numbers that are reported reflect full spreads (rather than half spreads). Therefore, the spread estimates reflect the average transaction costs for a round trip trade in the given universe of stocks. The standard deviations of the spread across stocks are also shown.

6 - Coverage in earlier periods is patchy. For example, bid and ask prices of stocks traded in NYSE are available only when the closing price is missing for the time period February, 1942 to December, 1992. Bid and ask information for NASDAQ stocks is available only for National Market Securities (NMS) from November 1982 to June 1992.

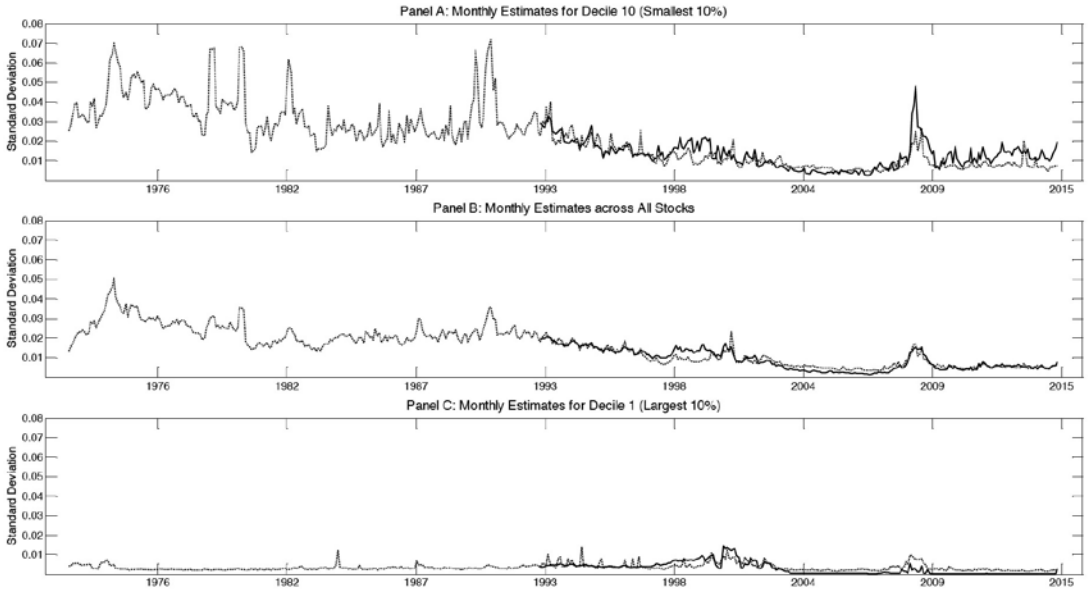
2. Descriptive Statistics on Transaction Costs

Exhibit 2: Effective Spread Estimates: Top 3,000 US stock universe
The figure shows the mean monthly spread estimates based on two estimators – the Range-Based Spread Estimator and the Closing Quoted Spread Estimator. Reported spreads are mean monthly 2-way spread estimates. Our sample universe consists of the 3,000 largest ordinary common stocks in the United States in each quarter based on market capitalisation. As the universe is re-sampled every quarter there may be more than 3,000 stocks in a given year. The daily spread estimate of each stock is estimated based on the chosen estimator. Monthly spreads of each stock are calculated as the average of daily spread estimates of those stocks with at least 12 days of daily spread estimates in a given month. The mean monthly spreads of top decile stocks (largest 10% of stocks), bottom decile stocks (smallest 10% of stocks) based on market capitalisation and the mean monthly spreads across all stocks in our sample universe are reported for each type of estimator. Range-Based spread estimates are estimated from January 1973 to December 2014, and due to limited data availability Closing Quoted spread estimates are estimated only from January 1993 to December 2014. Data Source: CRSP.

Plot A: Monthly Spread Estimates



Plot B: Standard Deviation of Monthly Spread



2. Descriptive Statistics on Transaction Costs

The results in Exhibit 2 suggest that both of our estimates provide overall similar results. This is true not only for the average cost estimate across stocks, but also for the 5th and 95th percentiles, which are similar across the two methods. The results also allow interesting conclusions to be drawn on the level, the time series variation, and the cross sectional variation of transaction costs. The average spread across all stocks had frequently reached values above 2% in the 1970s, but is situated clearly below 1% in the recent part of the sample. In such an equal-weighted average across 3,000 stocks, small stocks with high spreads obviously have a high influence. When looking at the top decile (i.e. the 300 largest stocks by market cap), the spread has taken on typical values in the area of 0.5% even during the early periods such as the 1970s. In contrast, the smallest decile stocks had historically reached spread levels exceeding 5%. We also observe spikes in the spread estimates which correspond to liquidity crisis. In particular, spikes are observed in the period from late 2008 to early 2009 – a period which saw major bank failures and a drying up of liquidity (Brunnermeier, 2009).

For the period since 1993, we can compare the two spread estimates that we consider. For the smallest decile stocks, we observe that both the range-based spread measure and the closing spread yield very similar results both in terms of level and in terms of time series variation. For the largest stocks, we also observe very similar time series behaviour with higher costs during the decline of technology stocks in the period 2000-2002 and the 2008-2009 banking crisis. The closing spread estimator, however, captures the decline in transaction costs after about 2003 in a more pronounced manner.

It is worth discussing how transaction costs behaved at points when market structure changed. In the US stock market, there are a few notable points when minimum tick sizes declined. The first change occurred in 1997 when the tick size was reduced from 1/8th to 1/16th and the second major reduction occurred in 2001 when the tick size went from 1/16th to 1/100th (i.e. or decimalisation). Smaller tick sizes allow for more competitive spreads. We can see that there is indeed a general reduction in spread levels if we compare the period prior to 1997 to the period after 2001.

2.2 Consistency with Transaction Cost Benchmarks

We evaluate our transaction cost estimates by comparing them with a variety of estimates from other sources or using other methodologies. It should be noted that cost estimates from low frequency data should not be judged by whether or not they precisely match the estimates from high frequency data. In fact, low frequency estimates have the advantage of much more general use compared to estimates from high frequency data. Such low frequency estimates will be useful as far as they are able to capture the main patterns in terms of time series variation and cross sectional differences in transaction cost levels. We first look at a variety of estimates for a broad sample of stocks which we either obtain in the literature or are able to compute ourselves. The most important benchmark for low frequency estimates used in the literature is a comparison with estimated spreads from high frequency databases such as TAQ. We also include other cost benchmarks based on plugging in proprietary cost records or using extrapolation.

2. Descriptive Statistics on Transaction Costs

However, all these comparisons remain comparatively imprecise because we only compare group-wise estimates of transaction costs. Moreover, the plug-in and extrapolation estimates are themselves questionable because they rely on particular models and model parameters. What we would be really interested in is to compare our stock level low frequency estimates of transaction costs to the corresponding high frequency measures. A stock-level comparison is more precise than a comparison of group-wise average transaction costs. Moreover, effective spreads calculated from high frequency data are a better benchmark than model-based estimates.

We are able to obtain direct estimates of effective spread from high frequency data from two sources. First, we obtain effective spreads computed by trading venues in compliance with Rule 605 regulation. This data is available from WRDS for a limited time period and with partial coverage of the broad stock universe. Second, we obtain effective spreads calculated from the TAQ intraday data for random selections of 300 stocks used in Hasbrouck (2009). Below, we describe each comparison in more detail and provide results.

2.2.1 Comparison with Group-wise Transaction Cost Levels

We first compare our estimates for different years and size categories to benchmark estimates. Chung and Zhang (2014) publish estimates for average spreads across all stocks and for each year, as well as estimates for different size categories based on effective spreads from high frequency data (TAQ). Below, we compare our estimates with their TAQ estimates. In order to ensure comparability, we ensure that the stock universe we use is comparable to that used in the Chung and Zhang (2014) analysis. Therefore, we deviate from our 3,000-stock universe created above, and instead follow Chung and Zhang's approach to constructing the universe. Their universe is based on all stocks available in CRSP and broken down into NYSE/AMEX and NASDAQ listed stocks. Moreover, we use two additional ways to obtain transaction costs from alternative methodologies and sources. First, we collect data from ITG's peer review reports. These reports contain quarterly post trade transaction cost estimates for different size categories of stocks. We match stocks in our universe to their ITG size category and apply the relevant cost estimate to these stocks for the comparison below. Second, we follow the Keim and Madhavan model, similar to that of Avramov, Chordia and Goyal (2006) because it is widely used in the literature. This model is based on the evidence in Keim and Madhavan (1997) linking transaction costs observed over a short time period to stock characteristics, and we follow Avramov, Chordia and Goyal in adding a time trend to these estimates. The different comparisons will be discussed in the first subsection below. A detailed description of the construction of the ITG and Keim and Madhavan estimates can be found in the appendix (available upon request). Table 1 shows results over time and across different size groups.

Our key benchmark for comparison is the TAQ based spread as it relies on a high quality database and corresponds to actual trade and quote data. We include the ITG estimates and Keim and Madhavan estimates for completeness. However, ITG estimates rely on a proprietary estimation methodology and thus cannot be replicated while the Keim and Madhavan estimates rely on extrapolation of a relationship between firm characteristics and transaction cost levels which were observed in a small dataset.

2. Descriptive Statistics on Transaction Costs

Table 1: Spread estimates across time from different sources and methodologies
The table shows the mean yearly spread estimates based on various estimators - Closing Quoted Spread Estimator, Corwin and Schultz Range-Based Spread Estimator, ITG peer review reports and Keim and Madhavan Estimator and compare those estimates with TAQ intra-day spread estimates reported by Chung and Zhang (2014). Reported spreads are mean yearly 2-way spread estimates. In order to compare our results with the results reported by Chung and Zhang (2014) we deviate from our sample universe of 3,000 largest stocks in United States based on market capitalisation to all the stocks in NYSE/AMEX and NASDAQ present in The CRSP database . The daily spread estimate of each stock is estimated for all estimators except that of the transaction cost measures extracted from ITG peer review reports. ITG reports quarterly transaction cost estimates (Implementation Shortfall = Delay Costs + Price Impact) for various size buckets such as large-cap, mid-cap and small-cap stocks which are then assigned to each stock based on the corresponding size bucket each stock falls into based on its market capitalisation and the yearly estimates are calculated by averaging across quarters. Yearly spreads of other daily estimates of each stock is calculated as the average of daily spread estimates of those stocks with at least 40 days of daily spread estimates in a given year. The yearly and quintile-based TAQ estimates are available for 1993 to 2009 from Chung and Zhang (2014). Data for the Closing Quoted Spread estimate is available only for the period from January 1993 to December 2014. Data for Range-Based Spreads and Keim and Madhavan estimates are available from January 1973 to December 2014, but the comparison analysis here is restricted to the period of CRSP Bid-Ask data availability (January 1993 to December 2014). Data from ITG's peer review reports are available from the third quarter of 2002 to the last quarter of 2014. The mean yearly estimates of all stocks in NYSE/AMEX and NASDAQ stocks on a year by year basis are reported in Panel A, which is the average of the yearly estimates across all stocks in a given year. The mean yearly estimates of all stocks in NYSE/AMEX and NASDAQ stocks for each size quintiles (market-cap quintiles) are reported in Panel B, which is the average of the yearly estimates across all stocks in a given market-cap quintile across all years from 1993 to 2009 (the longest available common time period across various estimates) except the ITG estimate, which is average of the yearly estimates across all stocks from the 3rd quarter of 2002 to 2014. Data Source: CRSP, ITG peer review reports available at <ww.ltg.com>, and Table 2 of Chung and Zhang (2014).

Panel A: Year by Year Transaction Cost Estimates

	NYSE/AMEX stocks						NASDAQ stocks				
Year	Spread	Closing Quoted Spread	Range-Based Spread	TAQ	ITG	Keim & Madhavan	Closing Quoted Spread	Range-Based Spread	TAQ	ITG	Keim & Madhavan
1993	Mean	3.13%	1.96%	1.75%		0.57%	8.25%	6.47%	5.09%		1.00%
	Median	**1.86%**	**1.11%**	**0.97%**		**0.00%**	**5.79%**	**4.53%**	**3.85%**		**0.74%**
1994	Mean	3.04%	1.80%	1.61%		0.54%	7.69%	5.99%	5.01%		1.04%
	Median	**1.89%**	**1.07%**	**0.97%**		**0.00%**	**5.57%**	**4.28%**	**3.84%**		**0.84%**
1995	Mean	3.42%	1.67%	1.51%		0.51%	7.07%	5.49%	4.63%		1.01%
	Median	**2.11%**	**0.97%**	**0.93%**		**0.00%**	**4.82%**	**3.91%**	**3.37%**		**0.73%**
1996	Mean	3.28%	1.50%	1.30%		0.44%	6.22%	5.01%	4.11%		0.95%
	Median	**2.07%**	**0.92%**	**0.81%**		**0.00%**	**4.44%**	**3.74%**	**3.05%**		**0.64%**
1997	Mean	3.62%	1.36%	1.11%		0.39%	5.46%	4.34%	3.83%		1.01%
	Median	**1.92%**	**0.81%**	**0.68%**		**0.00%**	**3.76%**	**3.08%**	**2.72%**		**0.71%**
1998	Mean	4.13%	1.37%	1.06%		0.43%	4.55%	3.85%	3.88%		1.08%
	Median	**2.15%**	**0.81%**	**0.65%**		**0.00%**	**3.20%**	**2.77%**	**2.70%**		**0.89%**
1999	Mean	4.75%	1.52%	1.29%		0.54%	3.58%	3.39%	3.19%		1.06%
	Median	**2.70%**	**0.94%**	**0.71%**		**0.00%**	**2.52%**	**2.58%**	**2.21%**		**0.82%**
2000	Mean	4.88%	1.80%	1.59%		0.62%	3.34%	3.56%	3.05%		1.06%
	Median	**2.59%**	**1.12%**	**0.77%**		**0.01%**	**2.42%**	**2.83%**	**2.20%**		**0.84%**
2001	Mean	3.31%	1.59%	1.40%		0.61%	3.58%	3.32%	3.30%		1.20%
	Median	**1.45%**	**0.85%**	**0.52%**		**0.00%**	**2.37%**	**2.35%**	**2.22%**		**1.08%**
2002	Mean	2.57%	1.65%	1.37%	3.12%	0.62%	3.04%	2.76%	3.10%	3.64%	1.23%
	Median	**1.13%**	**0.91%**	**0.44%**	**3.75%**	**0.00%**	**1.84%**	**1.98%**	**1.93%**	**3.75%**	**1.21%**
2003	Mean	1.52%	1.22%	0.96%	1.52%	0.59%	1.93%	1.97%	2.23%	1.73%	1.11%
	Median	**0.53%**	**0.69%**	**0.32%**	**1.80%**	**0.00%**	**1.08%**	**1.42%**	**1.42%**	**1.80%**	**0.92%**

2. Descriptive Statistics on Transaction Costs

2004	Mean	0.75%	0.93%	0.62%	1.45%	0.47%	1.18%	1.52%	1.30%	1.80%	0.92%
	Median	**0.20%**	**0.61%**	**0.23%**	**1.93%**	**0.00%**	**0.62%**	**1.29%**	**0.80%**	**1.93%**	**0.46%**
2005	Mean	0.66%	0.92%	0.65%	1.01%	0.47%	1.01%	1.33%	1.12%	1.23%	0.89%
	Median	**0.19%**	**0.61%**	**0.21%**	**0.71%**	**0.00%**	**0.47%**	**1.08%**	**0.60%**	**1.32%**	**0.36%**
2006	Mean	0.52%	0.86%	0.54%	0.99%	0.46%	0.78%	1.17%	0.85%	1.17%	0.82%
	Median	**0.14%**	**0.62%**	**0.17%**	**0.78%**	**0.00%**	**0.36%**	**0.99%**	**0.41%**	**1.25%**	**0.27%**
2007	Mean	0.60%	1.01%	0.56%	0.77%	0.45%	0.84%	1.21%	0.85%	0.91%	0.82%
	Median	**0.17%**	**0.74%**	**0.20%**	**0.59%**	**0.00%**	**0.36%**	**1.03%**	**0.35%**	**0.97%**	**0.32%**
2008	Mean	1.20%	1.92%	0.93%	3.74%	0.62%	2.54%	2.48%	1.80%	5.57%	1.17%
	Median	**0.25%**	**1.43%**	**0.37%**	**2.13%**	**0.05%**	**0.94%**	**2.04%**	**0.79%**	**8.27%**	**1.06%**
2009	Mean	0.90%	1.80%	1.02%	2.27%	0.74%	2.95%	2.58%	1.92%	3.31%	1.34%
	Median	**0.18%**	**1.27%**	**0.39%**	**2.15%**	**0.08%**	**0.87%**	**1.93%**	**0.86%**	**4.56%**	**1.46%**
2010	Mean	0.42%	1.03%		1.69%	0.57%	1.43%	1.50%		2.39%	1.15%
	Median	**0.08%**	**0.81%**		**1.18%**	**0.00%**	**0.40%**	**1.20%**		**2.28%**	**0.96%**
2011	Mean	0.41%	1.02%		1.68%	0.53%	1.23%	1.46%		2.35%	1.09%
	Median	**0.07%**	**0.81%**		**1.20%**	**0.00%**	**0.33%**	**1.18%**		**2.06%**	**0.82%**
2012	Mean	0.46%	0.92%		1.52%	0.53%	1.34%	1.35%		2.09%	1.05%
	Median	**0.07%**	**0.65%**		**1.19%**	**0.00%**	**0.33%**	**1.02%**		**1.97%**	**0.73%**
2013	Mean	0.37%	0.82%		1.27%	0.42%	0.98%	1.15%		1.63%	0.90%
	Median	**0.05%**	**0.59%**		**1.19%**	**0.00%**	**0.29%**	**0.92%**		**1.82%**	**0.36%**
2014	Mean	0.33%	0.83%		1.44%	0.38%	0.87%	1.21%		1.95%	0.81%
	Median	**0.10%**	**0.62%**		**1.23%**	**0.00%**	**0.27%**	**1.00%**		**2.22%**	**0.24%**

Panel B : Transaction Cost Estimates for Size Quintiles (1993 to 2009), ITG estimates are for 3rd Quarter 2002 to 2014)

	NYSE/AMEX stocks						NASDAQ Stocks				
Quintile	Spread	Closing Quoted Spread	Range-Based Spread	TAQ	ITG	Keim & Madhavan	Closing Quoted Spread	Range-Based Spread	TAQ	ITG	Keim & Madhavan
1	Mean	7.40%	3.47%	3.24%	2.78%	1.58%	9.71%	7.47%	6.81%	2.81%	1.68%
	Median	**4.84%**	**2.45%**	**2.13%**	**1.97%**	**0.00%**	**7.65%**	**5.74%**	**5.73%**	**2.37%**	**0.74%**
2	Mean	2.60%	1.40%	1.11%	2.23%	0.79%	4.88%	3.80%	3.97%	2.81%	1.51%
	Median	**2.03%**	**1.09%**	**0.90%**	**1.93%**	**0.00%**	**4.10%**	**3.27%**	**3.56%**	**2.22%**	**0.84%**
3	Mean	1.52%	0.98%	0.63%	1.54%	0.20%	3.18%	2.86%	2.48%	2.56%	1.17%
	Median	**1.30%**	**0.82%**	**0.53%**	**1.23%**	**0.00%**	**2.64%**	**2.48%**	**2.24%**	**1.97%**	**0.73%**
4	Mean	1.07%	0.82%	0.39%	1.18%	0.05%	2.07%	2.30%	1.63%	1.91%	0.65%
	Median	**0.99%**	**0.72%**	**0.31%**	**1.16%**	**0.00%**	**1.60%**	**2.02%**	**1.43%**	**1.82%**	**0.64%**
5	Mean	0.71%	0.73%	0.21%	0.90%	0.01%	1.07%	1.68%	0.86%	1.35%	0.19%
	Median	**0.66%**	**0.67%**	**0.16%**	**0.71%**	**0.00%**	**0.66%**	**1.52%**	**0.66%**	**1.19%**	**0.71%**

2. Descriptive Statistics on Transaction Costs

Analysing the year by year results in Panel A of Table 1, we can see that our two types of estimates match the estimates reported based on TAQ data reasonably well while our estimates appear conservative in the sense that they rather tend to indicate higher cost levels. For NYSE/AMEX stocks, the closing quoted spread tends to overestimate costs relative to both range-based estimates and the TAQ database until about 2000 to 2002. The closing quoted spreads however closely resemble the TAQ benchmark thereafter. For NASDAQ stocks, range-based and closing quoted spread estimates resemble the TAQ benchmark. From about 2002 onwards, the closing quoted spread tracks the TAQ benchmark more closely than the range-based spread estimates. The ITG peer review data tends to suggest larger spreads while the Keim and Madhavan estimator suggests much lower spreads than our two methods.

When looking at the size categories and focusing on the median values, we again observe that our estimates seem close to the TAQ based estimates but tend to include a slight upward bias indicating that any conclusions from applying these estimates to assessing the viability of investment strategies after trading costs will be rather conservative. In particular, the closing quoted and range-based spread estimators both tend to overestimate the magnitude of spreads compared to the TAQ benchmark, especially for NYSE stocks. The range-based spread estimator overestimates to a lesser degree, especially for the smaller stocks. Moreover, the closing quoted and range-based spread estimators deliver results which are quite in line with the ITG estimates, especially in the four largest size quintiles. Again, the Keim and Madhavan estimates give much lower magnitude than all other methods.

Overall, we conclude that closing spread and range-based spread reliably capture levels of high frequency (TAQ) transaction costs during different time periods and for different size categories, while providing a somewhat conservative view.

After such group-wise comparisons, we now turn to two comparisons with high frequency transaction cost benchmarks at the stock level.

2.2.2 Comparison with Rule 605 Data

We have access to data from Rule 605 filings of trading venues. This data contains monthly stock level estimates of transaction costs based on trade execution data. This dataset is a result of regulation, specifically SEC-mandated disclosure about execution quality of stock trades (Rule 605 is an evolution of the former Rule 11Ac1-5).

The data includes, among other things, information about the number of shares executed in several time intervals after the order was placed, information on shares executed outside of the National Best Bid and Offer, information about price improvement on the trades and more importantly, the dataset contains information about average realised and effective spreads.

However, Rule 605 data is only available to us for a limited time period and for a partial universe of stocks, as published by WRDS. WRDS compiles the data for the period from June 2001 to August 2005. For quarterly reporting, we only use quarters that have data for all 3 months and we only use stocks for which there are matching identifiers with our database sourced from CRSP. This execution quality information is available from WRDS for each stock monthly. For each stock and

2. Descriptive Statistics on Transaction Costs

each month, the data is further broken down into four order size categories based on number of shares in a given order (100-499 shares, 500-1,999 shares, 2,000-4,999 shares and 5,000-9,999 shares), five different order types (market, marketable limit, inside the quote limit, at the quote limit, near the quote limit) and across six different trade venues: AMEX (American Stock Exchange), BSE (Boston Stock Exchange), CHX (Chicago Stock Exchange), CSE (Cincinnati Stock Exchange), NYSE (New York Stock Exchange) and PHLX (Philadelphia Stock Exchange).

We use the effective spread as reported in the data. The effective spread is twice the distance of the trade price from the mid-point at the time of the trade. To obtain our estimate of the Rule 605 effective spread we compute the weighted average of the reported effective spread using the share volume of the trade size buckets as the weights. We do this for the market and marketable limit orders for a particular stock each month depending on the data availability.

We conduct a stock level comparison between the WRDS data and our estimates, which is discussed in the second subsection below. We note that Rule 605 data is a common benchmark used in the empirical literature against which to assess the quality of trading cost estimates (O'Hara and Ye, 2011; Goyenko, Holden and Trzcinka, 2009; Hasbrouck, 2007; Zhao and Chung, 2007; Bennett and Wei, 2006; Boehmer, 2005).

The subsection also assesses the information content of quarterly mean Range-Based and Closing Quoted spread estimates with respect to quarterly spread estimates from Rule 605 data. The following cross-sectional regression has been run for each quarter from July, 2001 to June, 2005:

$$S_{i,q}^{R605} = \alpha_q + \beta_q * S_{i,q}^{EST} + \varepsilon_{i,q}$$

where $S_{i,q}^{EST}$ is the quarterly average spread estimate and $S_{i,q}^{R605}$ is quarterly average Rule 605 spread for stock *i* in quarter *q*. OLS estimates are reported in Table 2.

This framework is commonly used in assessing the information content of proxies for a given indicator (Christensen and Prabhala, 1998). If the estimated spread is an unbiased estimate of the true spread measure, then α_q =1 and β_q =1. However, if β_q is close to 1 and α_q is significantly negative, then we can conclude that our spread estimate is more conservative than the true spread as it may also capture price impact along with the spread.

Table 2 indicates the different measures of comparison of our estimates with the cost levels indicated in Rule 605 reporting.

The results suggest a strong link of our cost estimates with the Rule 605 data. In terms of correlation, transaction costs estimated using the high and low or the closing spreads matches the information from Rule 605 reports relatively well. However, the average differences show that our estimates are larger than the average Rule 605 costs. Overall, these results confirm that our estimates reliably capture the structure of costs in the cross section and in the time series of stocks, while providing somewhat conservative estimates.

2. Descriptive Statistics on Transaction Costs

Table 2: Comparison of Range-Based and Closing quoted spreads with Rule 605 spreads
The table compares Range-Based and Closing quoted spread estimates with effective spreads as reported through a SEC-mandated disclosure (Rule 605, formerly Rule 11Ac1-5). The effective spread is twice the distance of the trade price from the mid-point at the time of the trade. The spread is reported monthly for each stock and for different trade size buckets across six different trade venues. To obtain our estimate of the Rule 605 effective spread we compute the weighted average of the reported effective spread using the share volume of the trade size buckets as the weights. We do this for the market and marketable limit orders for a particular stock each month depending on the data availability. The source of the data is WRDS that compiles the data for six trading venues for the period from June 2001 to August 2005. For quarterly reporting, we only use quarters that have data for all 3 months and we only use stocks for which there are matching identifiers with our database sourced from CRSP. The number of stocks that are included in the comparison is indicated for each quarter. Note that data availability for the Rule 605 data is rather limited. Only a small percentage of stocks in our Total Market Universe have Rule 605 data. We compute quarterly average spread by averaging across the months. For the estimated spreads (Closing Quoted and Range-Based) that come with daily frequency, we require at least 40 daily observations in the quarter to obtain the quarterly average. For the estimates of Rule 605 spreads that come with monthly frequency we average across the 3 monthly estimates. OLS estimates for the regression to assess the information content of Range-Based and Closing Quoted spread estimates, along with correlation, root mean squared error and difference between average spreads, are reported in Panel A and B respectively. Mean difference is computed by averaging spread estimates across all stocks and subtracting average Rule 605 spread from the respective average spread estimate. Relative mean difference simply divides the latter by average Rule 605 spread. The root mean squared error reported below is based on the difference between Rule 605 and low-frequency spread estimates and NOT on the errors of regression. Coefficients that are statistically significant at the 95% level are reported in bold. Data Source: CRSP, WRDS.

Panel A: Comparison of Range-Based Spread estimates with Rule 605 spreads

Quarter	Number of Stocks	Regression Intercept	Regression Slope	Adjusted R^2	Correlation	RMSE	Mean Difference	Relative Mean Difference
2001-Q3	1164	**-0.45%**	**1.32**	58.79%	0.77	0.87%	0.15%	19.74%
2001-Q4	1168	**-0.37%**	**1.17**	53.88%	0.73	0.97%	0.21%	26.15%
2002-Q1	1301	**-0.15%**	**0.65**	32.28%	0.57	0.82%	0.49%	101.00%
2002-Q2	1385	**-0.21%**	**0.65**	38.99%	0.62	0.88%	0.57%	123.58%
2002-Q3	1266	**-0.21%**	**0.57**	38.62%	0.62	1.09%	0.76%	147.52%
2002-Q4	1299	**-0.22%**	**0.61**	39.52%	0.63	0.98%	0.68%	133.65%
2003-Q1	1292	**-0.31%**	**0.76**	31.34%	0.56	0.95%	0.54%	127.52%
2003-Q2	1243	**-0.70%**	**1.30**	35.81%	0.60	1.20%	0.44%	114.72%
2003-Q3	1293	**-0.24%**	**0.72**	47.41%	0.69	0.61%	0.45%	149.57%
2003-Q4	1307	**-0.16%**	**0.63**	42.52%	0.65	0.55%	0.41%	160.13%
2004-Q1	1283	**-0.05%**	**0.43**	30.38%	0.55	0.62%	0.46%	183.54%
2004-Q2	1295	**-0.20%**	**0.64**	43.56%	0.66	0.60%	0.46%	179.71%
2004-Q3	1295	**-0.20%**	**0.66**	44.82%	0.67	0.59%	0.44%	168.09%
2004-Q4	1307	**-0.23%**	**0.71**	52.61%	0.73	0.54%	0.43%	180.42%
2005-Q1	1335	**-0.15%**	**0.54**	41.85%	0.65	0.58%	0.46%	210.27%
2005-Q2	1376	**-0.18%**	**0.59**	43.25%	0.66	0.58%	0.46%	202.53%

Panel B: Comparison of Closing Quoted Spread estimates with Rule 605 spreads

Quarter	Number of Stocks	Regression Intercept	Regression Slope	Adjusted R^2	Correlation	RMSE	Mean Difference	Relative Mean Difference
2001-Q3	1164	**-0.18%**	**0.71**	78.19%	0.88	0.95%	0.58%	74.64%
2001-Q4	1168	**-0.18%**	**0.61**	64.51%	0.80	1.36%	0.81%	102.02%
2002-Q1	1301	**0.03%**	**0.33**	33.08%	0.58	1.40%	0.91%	185.30%
2002-Q2	1385	**-0.10%**	**0.66**	69.88%	0.84	0.68%	0.39%	85.39%
2002-Q3	1266	**-0.08%**	**0.56**	70.97%	0.84	0.92%	0.56%	107.64%
2002-Q4	1299	**-0.09%**	**0.62**	66.85%	0.82	0.76%	0.44%	87.54%
2003-Q1	1292	**-0.15%**	**0.67**	48.98%	0.70	0.84%	0.43%	101.62%
2003-Q2	1243	**-0.18%**	**0.85**	28.23%	0.53	1.20%	0.28%	71.71%
2003-Q3	1293	-0.01%	**0.78**	83.22%	0.91	0.27%	0.10%	32.83%

2. Descriptive Statistics on Transaction Costs

2003-Q4	1307	**0.10%**	**0.56**	50.90%	0.71	0.38%	0.03%	11.84%
2004-Q1	1283	**0.05%**	**0.76**	74.79%	0.86	0.21%	0.01%	5.41%
2004-Q2	1295	**0.02%**	**0.84**	75.80%	0.87	0.24%	0.02%	8.01%
2004-Q3	1295	0.00%	**0.91**	84.81%	0.92	0.19%	0.03%	10.25%
2004-Q4	1307	-0.01%	**0.91**	86.50%	0.93	0.17%	0.03%	13.17%
2005-Q1	1335	**-0.04%**	**1.00**	86.08%	0.93	0.14%	0.04%	16.49%
2005-Q2	1376	**-0.03%**	**0.94**	90.14%	0.95	0.14%	0.04%	18.28%

2.2.3 Comparison with TAQ Data

This subsection compares yearly Range-Based and Closing Quoted spread estimates to the yearly spreads obtained from Trade and Quote database. The comparison is done on yearly basis and cross-sectional correlation, root mean squared error and difference between average spreads are reported for each year in Table 3. Besides that, we also assess information content of Range-Based and Closing Quoted spread estimates by running following regression for each year from 1993 to 2005:

$$S_{i,y}^{TAQ} = \alpha_y + \beta_y * S_{i,y}^{EST} + \varepsilon_{i,y}$$

where $S_{i,y}^{TAQ}$ corresponds to yearly spread estimate from TAQ database and $S_{i,y}^{EST}$ corresponds to yearly spread estimate from low frequency data for stock *i* in year *y*. The comparison sample consists of approximately 150 NASDAQ and 150 NYSE/Amex randomly selected firms per each year.

7 - Available at : <http://people.stern.nyu.edu/jhasbrou/Research/GibbsEstimates2006/Liquidity%20estimates%202006.htm>

Table 3: Comparison of Range-Based and Closing quoted spreads with TAQ spreads
The table compares yearly mean spread estimates to the yearly mean effective spreads reported by Trade and Quote (TAQ) from 1993 to 2005. Yearly TAQ spreads for approximately 300 capitalisation-stratified randomly selected stocks are provided by Joel Hasbrouck on NYU Stern web page.[7] Yearly estimates for Range-Based and Closing quoted spreads are computed from daily spread estimates with at least 40 daily observations per year. OLS estimates for the regression to assess the information content of Range-Based and Closing Quoted spread estimates, along with correlations, root mean squared error and difference between average spreads, are reported in Panel A and Panel B respectively. Mean difference is computed by averaging spread estimates across all stocks and subtracting average TAQ spread from the respective average spread estimate. Relative mean difference simply divides the latter by average TAQ spread. The root mean squared error reported below is based on the difference between TAQ and low-frequency spread estimates and NOT on the errors of regression. Coefficients that are statistically significant at the 95% level are reported in bold. Data source: CRSP, Hasbrouck's web page.

Panel A: Comparison of Range-Based Spread estimates with TAQ spreads

Quarter	Number of Stocks	Regression Intercept	Regression Slope	Adjusted R^2	Correlation	RMSE	Mean Difference	Relative Mean Difference
1993	289	-0.08%	**0.98**	86.23%	0.93	1.83%	0.15%	4.71%
1994	271	**-0.69%**	**1.17**	85.97%	0.93	1.75%	0.13%	3.97%
1995	288	0.18%	**0.92**	82.83%	0.91	2.02%	0.06%	2.11%
1996	290	**-0.35%**	**1.11**	85.07%	0.92	1.26%	0.05%	1.78%
1997	290	-0.10%	**0.97**	82.21%	0.91	1.27%	0.17%	6.94%
1998	293	**0.05%**	**0.93**	73.38%	0.86	1.21%	0.11%	5.42%
1999	292	**-0.24%**	**1.02**	72.13%	0.85	1.30%	0.20%	9.60%
2000	294	-0.18%	**0.91**	69.52%	0.83	1.46%	0.40%	18.62%
2001	296	-0.22%	**1.12**	75.58%	0.87	1.56%	-0.04%	-1.77%
2002	292	**-0.52%**	**1.25**	67.18%	0.82	1.35%	0.06%	3.48%
2003	294	**-0.35%**	**1.06**	77.11%	0.88	0.98%	0.27%	22.38%
2004	292	**-0.30%**	**1.05**	57.16%	0.76	0.96%	0.25%	26.36%
2005	295	**-0.42%**	**1.12**	52.91%	0.73	0.85%	0.29%	36.94%

2. Descriptive Statistics on Transaction Costs

Panel B: Comparison of Closing Quoted Spread estimates with TAQ spreads

Quarter	Number of Stocks	Regression Intercept	Regression Slope	Adjusted R^2	Correlation	RMSE	Mean Difference	Relative Mean Difference
1993	288	**-0.41%**	**0.79**	97.02%	0.99	2.17%	1.39%	42.72%
1994	271	**-0.56%**	**0.81**	96.80%	0.98	2.15%	1.44%	44.78%
1995	288	**-0.61%**	**0.81**	94.69%	0.97	2.17%	1.45%	48.74%
1996	290	**-0.47%**	**0.76**	91.78%	0.96	2.17%	1.50%	54.51%
1997	289	**0.24%**	**0.52**	72.19%	0.85	3.20%	1.80%	72.52%
1998	293	**0.80%**	**0.34**	41.26%	0.64	4.85%	1.68%	81.56%
1999	292	**1.03%**	**0.23**	53.10%	0.73	4.93%	2.37%	114.80%
2000	294	**0.78%**	**0.34**	65.87%	0.81	3.61%	1.86%	87.29%
2001	294	-0.07%	**0.70**	92.45%	0.96	1.74%	1.05%	46.56%
2002	292	-0.10%	**0.82**	93.89%	0.97	0.88%	0.52%	29.69%
2003	294	-0.01%	**0.87**	95.32%	0.98	0.47%	0.20%	16.49%
2004	292	0.05%	**0.98**	97.16%	0.99	0.24%	-0.03%	-3.68%
2005	295	0.02%	**1.06**	97.72%	0.99	0.20%	-0.06%	-8.09%

The results suggest that both estimators achieve high R-squared most of the time, while the closing quoted spread dominates the range-based spread except for the 1997-2000 period. The range-based spread has negative alpha (whenever it is significant) and beta less than one for most of the time periods. This implies that estimates are conservative. Except for the years from 1997 to 2000, the closing quoted spread achieves higher correlation and R-squared with the benchmark compared to the range-based spread. Negative alphas and betas less than one indicate that it is even more conservative than range-based spreads. For the years 1997-2001, the closing quoted spread is dominated by the range-based spread in capturing the true spread (lower R-squared).

Overall, our empirical comparison with cost estimates resulting from high frequency data suggests that our estimates capture the cross sectional and time series structure of transaction costs, while providing somewhat conservative estimates.[8]

8 - It should also be noted that our analysis, when based on volume data, actually tends to underestimate the available trading volume given that additional volume may be unreported to data sources (e.g. in the case of dark pools, internal crossing platforms, etc.). In this sense, any estimate of implementation issues may be biased upwards suggesting that our analysis provides a conservative conclusion on the implementability of smart beta strategies.

2.3 Relationship with Market Conditions

This section looks at the relationship of our spread estimates with exogenous variables which have not been used to generate the effective spread estimates. Intuitively, one would expect that the time variation of average transaction costs is related to market variables such as volatility or measures of market illiquidity.

We use two suitable measures of market states that have been shown to be associated with illiquidity. First, the VIX which reflects option-implied volatility has been shown to be related to illiquidity as the option implied volatility spikes when expected volatility and/or risk aversion increase(s), which is often the case in periods of illiquidity. The VIX is not a direct measure of liquidity but highly related to market liquidity and thus has been used as an empirical proxy for market liquidity (Nagel, 2012). Another measure is the TED spread, which is simply the difference between the interbank lending rate and the treasuries rate at an identical short maturity (typically 3 months). Frazzini and Pedersen (2014)

2. Descriptive Statistics on Transaction Costs

and Garleanu and Pedersen (2011) see this variable as indicating funding constraints (funding liquidity risk). It could thus be argued that the VIX and the TED spread may capture different aspects of liquidity, in particular market liquidity and funding liquidity.

In order to assess how our spread estimates are related to state variables reflecting liquidity conditions, we run the following regression.

The regression uses monthly change in spread estimates and monthly changes in VIX and TED spread. The *VIX is scaled by* $1/\sqrt{12}$ *to obtain monthly measure.*

$$\Delta Spread_t = \alpha + \beta^{VIX} \Delta VIX_t + \beta^{TED} \Delta TEDSpread_t + \varepsilon_t$$

where $\Delta Spread_t$ corresponds to the change in either Closing Quoted Spread or the change in Range-Based spread, α is the intercept, β^{VIX} and β^{TED} are the regression coefficients corresponding to the independent variables and ε_t is the error term.

We are interested in particular in the slope coefficients with respect to the VIX and TED spread. The resulting regression statistics are shown in Table 4. We also provide the average spread measures conditional upon the VIX and TED Spread.

The results in Table 4 suggest that our spread proxies are positively related to state variables capturing illiquidity. We can thus be confident that applying such

Table 4: Relationship with Market State Variables
For the spread estimates (Closing Quoted and Range-Based Spreads) we require at least 12 daily observations for a single security to get a monthly average. We then average spread measures across the 3,000-stock universe used in our analysis. The final monthly estimate is then a cross-sectional mean of the estimated individual monthly spreads. VIX data comes from Bloomberg and the VIX is scaled by $1/\sqrt{12}$ *to obtain monthly measure. TED spread data come from St Louis FED. The analysis starts in 1993 due to data availability of Closing Quoted Spread. Newey-West robust t-statistics are used to compute p-values. Coefficients significant at the 95% level are highlighted in bold. Data Source: CRSP, Bloomberg, St Louis FED.*

PANEL A: Regression
Regressions use monthly change in spread estimates and monthly change in VIX and TED spread. The regressions below are estimated using monthly data (more precisely, the changes of monthly values).

Full period: January 1993 to December 2014						
	Change Closing Quoted Spread			Change in Range-Based Spread		
Intercept	0.00	0.00	0.00	0.00	0.00	0.00
p-value	(0.10)	(0.13)	(0.10)	(0.52)	(0.61)	(0.52)
Change in VIX	**0.02**		**0.01**	**0.04**		**0.04**
p-value	(0.00)		(0.00)	(0.00)		(0.00)
Change in TED spread		**0.09**	**0.07**		**0.12**	0.08
p-value		(0.00)	(0.01)		(0.05)	(0.17)
R-squared	4.72%	4.21%	7.44%	7.50%	2.29%	8.47%

Recent period: January 2001 to December 2014						
	Change Closing Quoted Spread			Change in Range-Based Spread		
Intercept	0.00	0.00	0.00	0.00	0.00	0.00
p-value	(0.24)	(0.27)	(0.24)	(0.64)	(0.70)	(0.64)
Change in VIX	**0.02**		**0.01**	**0.05**		**0.04**
p-value	(0.01)		(0.02)	(0.01)		(0.02)
Change in TED spread		**0.11**	**0.09**		**0.21**	**0.15**
p-value		(0.00)	(0.00)		(0.00)	(0.03)
R-squared	9.35%	12.66%	17.55%	10.75%	7.34%	14.47%

2. Descriptive Statistics on Transaction Costs

PANEL B: Conditional Mean

Conditional Mean is the average spread calculated conditional upon the VIX and TED spread variables. Classification is done on monthly data with the top 50% months based on the conditional variables classified as a low liquidity regime and the bottom 50% based on the conditional variables classified as a high liquidity regime. VIX is scaled by $1/\sqrt{12}$ to obtain monthly measure.

Full period: January 1993 to December 2014	Monthly Implied Volatility (VIX)	Closing Quoted Spread	Range-Based Spread
Average across Top 50% Months based on Monthly Implied Volatility (Low Liquidity Regime)	7.52%	0.99%	1.31%
Average across Bottom 50% Months based on Monthly Implied Volatility (High Liquidity Regime)	4.13%	0.94%	1.15%

	TED Spread	Closing Quoted Spread	Range-Based Spread
Average across Top 50% Months based on TED Spread (Low Liquidity Regime)	0.76%	1.30%	1.39%
Average across Bottom 50% Months based on TED Spread (High Liquidity Regime)	0.25%	0.63%	1.08%

Recent period: January 2001 to December 2014	Monthly Implied Volatility (VIX)	Closing Quoted Spread	Range-Based Spread
Average across Top 50% Months based on Monthly Implied Volatility (Low Liquidity Regime)	7.77%	0.55%	1.25%
Average across Bottom 50% Months based on Monthly Implied Volatility (High Liquidity Regime)	4.21%	0.24%	0.81%

	TED Spread	Closing Quoted Spread	Range-Based Spread
Average across Top 50% Months based on TED Spread (Low Liquidity Regime)	0.68%	0.41%	1.10%
Average across Bottom 50% Months based on TED Spread (High Liquidity Regime)	0.20%	0.38%	0.95%

spread estimates to smart beta strategies will not only properly account for cross sectional differences in transaction costs but will also penalise strategies which trade heavily in market states where transaction costs are high because of low liquidity in the market. After having established that our two spread estimates behave in an expected way and are well suited to reliably capture transaction costs, we now turn to discussing how we are going to apply cost measures to smart beta strategies.

2.4 Retaining a Suitable Method

We follow Giglio and Shue (2014) in drawing on the Chung and Zhang (2014) estimate of transaction costs which uses the closing spread. We prefer this measure because of its superior empirical performance, as discussed above. However, since this measure is only available from 1993 due to limited data availability, we use the range-based spread estimator for the time period prior to 1993. The advantage of this approach is that we use the most reliable spread estimator while having the possibility to extend the analysis back in time. However, our results are not likely to depend heavily on this particular choice. As discussed above, both estimators provide broadly similar results in capturing information on transaction costs. Moreover, we have conducted further analysis showing the similarity of the transaction cost estimates obtained with the closing spread estimator compared to the range-based estimator over the period when both are available. This analysis is available in the separate appendix (available upon request).

3. Analysing Smart Beta Strategies

3. Analysing Smart Beta Strategies

We analyse a range of strategies including generic strategies applied to US stocks, as well as a set of available equity strategy indices. The reason for analysing generic strategies is that we can control parameters such as the breadth of the stock universe explicitly and thus analyse the dependence of transaction costs and implementation measures on such parameters. The advantage of using available indices where one cannot control such parameters is that we can draw conclusions that are directly applicable to investable indices that are available for use by investors.

For each of the strategies, along with the cap-weighted benchmark, the following three types of measures are reported:

- **Implementation Costs** – Measures such as Turnover, Transaction costs and Days to Trade that measure the costs associated with implementing various smart beta strategies are grouped under implementation costs;
- **Performance Measures** – General performance measures such as Returns, Relative Returns, Sharpe Ratio and Information Ratio for both gross and net return series of the smart beta indices are reported under performance measures;
- **Liquidity Measures** – Liquidity measures help in assessing the liquidity of the smart beta strategies in relation to the cap-weighted reference. Market capitalisation, spread and Amihud liquidity measures are used as parameters to assess the liquidity of the strategies.

Before discussing the results, we now provide a detailed definition of the measures in each category and how they are computed, before turning to the results.

Measures of implementation cost

First, in terms of implementation costs and related measures, we retain the following indicators:

One-Way Turnover

One-way Turnover is half the sum of absolute differences of individual weights between the end of the previous rebalancing quarter and the beginning of the current rebalancing quarter. One-way Turnover is annualised for each quarter by multiplying it by 4.

$$One - Way\ Turnover_q = \frac{\sum_{i=1}^{N} abs(w_{q,i} - w_{q-1,i})}{2}$$

where N is the number of stocks in both the quarters combined, q is the current quarter, $w_{q,i}$ is the weight of the stock i in quarter q.

Transaction Costs

The transaction cost estimates use the spread estimates according to the year of the rebalancing – Range-Based spread until 1993 and Closing Quoted Spread from 1993 onwards. The reported transaction cost estimates are the difference between the annualised gross and net returns. Net returns are obtained after accounting for transaction costs at each quarterly rebalancing by multiplying the change in weight of each stock (including stock deletions and additions) between the final weight before rebalancing and the optimal weights after rebalancing by half spreads.

Days to Trade (95th %ile)

Days to trade (DTT) refers to the time it takes to trade an initial investment given the stock weights at every quarterly rebalancing date, assuming 10% of the average daily dollar traded volume can be traded. The nominal amount considered for the initial investment equals USD 1bn and is deflated back in time in line with the return of the cap-weighted market index.

3. Analysing Smart Beta Strategies

The average 95th percentile of DTT is reported, which is the time series average of the cross-sectional 95th percentile of DTT at each quarterly rebalancing.

$$DTT_{i,q} = Notional\ Amount_q * \frac{w_{i,q}}{0.10 * Average\ Daily\ Traded\ Volume_{i,q}}$$

where $DTT_{i,q}$ is the days to trade of stock i in quarter q, $Notional\ Amount_q$ is the deflated investment of USD 1bn according to the return of the cap-weighted market index in quarter q and $w_{i,q}$ is the weight of the stock i in quarter q. $Average\ Daily\ Traded\ Volume_{i,q}$ is the average traded volume of stock i in quarter q and it is multiplied by the assumed percentage of tradability.

Measures of performance

Second, in terms of performance measures, we provide measures of both gross and net return series of the smart beta indices. All statistics are annualised. Mean returns are based on geometric mean. The gross return cap-weighted index is used as the benchmark.

Measures of liquidity

Third, in terms of liquidity measures, for each liquidity measure both the average of the weighted average measure in each quarter and the average extreme measure is reported. Extreme liquidity measures are liquidity measures of the most difficult to trade stocks in the smart beta strategies. This gives an overview of a worst case scenario as opposed to an average case scenario enabling investors to thoroughly assess the investability of a particular smart beta strategy. Extreme liquidity measures are obtained by sorting the universe every quarter by the corresponding liquidity measure and the measure corresponding to the stock that represents the 95% cumulative weight in the corresponding smart beta strategy is reported as an extreme measure in each quarter. For example, extreme market capitalisation represents the market capitalisation of the smallest stock that represents the 95% cumulative weight in the portfolio when stocks and their weights are sorted from largest market capitalisation to smallest. The average market capitalisation of an index is the weighted averaged market capitalisation of the index.

Amihud Measure

The illiquidity measure developed by Amihud (2002) is one of the most widely used liquidity proxies in the finance literature to measure price impact of trading. The Amihud Measure of a stock is calculated every day by dividing the absolute return for the day by its volume traded for the day. For each quarter, the average Amihud measure of each stock in that quarter is considered for weighted average and extreme measure.

$$Amihud\ Measure_{i,q} = \frac{\sum_{n=1}^{N}\left(\frac{abs(ret_{i,n})}{vol_{i,n}}\right)}{N}$$

where N is the number of days in quarter q, $ret_{i,n}$ and $vol_{i,n}$ are the return and volume traded of stock i in day n.

Spread

Spread is the two-way spread either based on the range-based estimator if the period of computation is before 1993 or based on the closing quoted spread estimator if the period of computation is 1993 or later.

It should be noted that our estimates of transaction costs capture the cost of periodic rebalancing in the strategies. This element of rebalancing of course constitutes the key difference with respect to cap-weighted indices which are close to buy and hold. Indeed, the main element

3. Analysing Smart Beta Strategies

of weight change in cap-weighted indices, is the reconstitution of index members, while smart beta indices are impacted both by reconstitution and by rebalancing of the weights back to target weights that correspond to the weighting strategy. Given that many investors may consider smart beta strategies as a replacement for cap-weighted indices, it may be relevant to analyse the switching cost as well. However, for long term investors in smart beta strategies the key cost component will be the costs of rebalancing while the switching costs are one-off costs faced initially. Therefore, we focus here on the costs of staying invested in such a strategies and we relegate results on switching costs to the appendix (available upon request).

We now turn to the results obtained with generic strategies before looking at investable indices in the following subsection.

3.1 Generic Strategies

We test the performance of the two alternative weighting schemes other than the traditional market-cap weighting, namely Equal Weighting and Fundamental Weighting. These two strategies are chosen as they consistently feature in the discussion on transaction costs and liquidity-related issues. We assess the common claims about transaction costs and liquidity made regarding these strategies for different universe sizes ranging from 250 stocks to 3,000 stocks (selected by the largest market capitalisation). This helps us assess how different strategies behave in terms of performance and ease of implementation in different stock universes, such as large-cap stock universes or broad equity market universes.

Based on the different universe definitions, we randomly select 100 stocks to create the equal-weighted or fundamentally weighted portfolios. This random selection process is repeated 1,000 times, meaning that for every combination of universe size and generic weighting scheme 1,000 generic indices are produced. This allows us to assess a range of outcomes for each weighting scheme and universe size chosen.

Our data sample consists of the 3,000 largest US ordinary common shares (i.e. shares with share code 10 or 11) from the CRSP database listed on the NYSE, Amex and NASDAQ stock exchanges. The 3,000 largest stocks are selected on every rebalancing day based on the stocks' total market capitalisation as of the previous day. All indices computed using alternative generic weighting schemes discussed in this section are rebalanced quarterly on the third Friday March, June, September and December every year. The data sample is extended as far back as possible where at least 3,000 stocks are available in the three stock exchanges combined. Thus, our sample period is from December 1972 to December 2014 (42 years). From the 3,000-stock database we test the generic strategies in different universe sizes such as 250, 500, 1,000 or 3,000 largest stocks respectively. In each scenario, for the chosen universe size, we randomly select 100 stocks at inception and apply the generic weighting scheme. These generic strategies are not subjected to any turnover or liquidity controls. On every subsequent quarterly rebalancing, we replace each stock that disappears from the restricted universe with another randomly chosen stock.

The fundamental-weighted index is constructed using four fundamental

3. Analysing Smart Beta Strategies

accounting data obtained from the Compustat database. They include book value (ceq), sales (sale), earnings (ni) and cash earnings (oibdp) with respective Compustat item names in brackets. The fundamental weight of each stock is calculated as follows: each stock is assigned four different weights, each of them proportional to the average of previous three year sales, earnings and cash earnings, as well as book value reported in previous year. From June of year y to March of year y+1, the fundamental data from year y-1, y-2 and y-3 are considered. The final weight is computed as an average of four fundamental weights. In case the average of any fundamental is negative, its respective weight is set to zero. If the book value is missing for a company, it is replaced by its market capitalisation. If an earnings weight is missing, weight of book value is assigned to it. If sales weight is missing, the average of previous two is assigned to it. If the weight of cash earnings is missing, the average of previous three is assigned to it. In case all weights are unavailable for a company, its final weight is determined as market capitalisation weight. If all fundamental weights are zero, then one fourth of market-cap weight is taken. The weights of other securities are renormalised accordingly if any of the above-mentioned cases occurs. The weights of those companies that have more than one class of shares in our universe are split in proportion to the market capitalisation of each class of share.

Table 5 shows the implementation metrics.

An inspection of the results in Table 5 suggests that, as the universe size increases, turnover, days to trade and transaction costs increase considerably. When the universe size only contains large-cap stocks (e.g. 250 or 500), equal weighting has similar transaction cost levels as those of fundamental weighting. However, as the universe size increases to include small and micro cap stocks, equal weighting has considerably higher transaction costs compared to those of fundamental weighted indices.

Table 6 provides additional metrics concerning performance measures.

3. Analysing Smart Beta Strategies

Table 5: Implementation Costs of Generic Alternative Weighting Schemes (USA Long Term Track Records (LTTR) – Long Term – 42 Years)
The time period of analysis is 31-Dec-1972 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. The generic alternative weighting schemes assessed in addition to the traditional market cap-weighting are equal weighting and fundamental weighting. Fundamental Weighting weights stocks based on fundamental variables such as Book Value, Sales, Earnings and Cash Earnings. Generic alternative weighted schemes are tested across various size of the underlying universe. From the 3,000 largest stocks in the USA, universes comprising the 250, 500, 1,000 and 3,000 largest stocks are chosen and from each universe 1,000 random samples of 100 stocks are selected and weighted according the generic weighting scheme chosen. For each combination of universe size and the weighting scheme chosen from the 1,000 portfolios formed, the implementation costs that correspond to 5th percentile, average and 95th percentile are reported here. One-way Turnover is half the sum of absolute deviations of individual weights between the end of a quarter and the beginning of the following quarter. One-way Turnover is annualised for each quarter by multiplying it by 4. The transaction cost estimates use the spread estimates according to the year of the rebalancing – Range-Based spread until 1993 and Closing Quoted Spread from 1993 onwards. The reported transaction cost estimates are the difference between the annualised gross and net returns. Net returns are obtained after accounting for transaction costs at each quarterly rebalancing by multiplying the change in weight of each stock between the final weight before rebalancing and the optimal weights after rebalancing, including stock deletions and additions. Days to trade (DTT) refer to the time it takes to trade an initial investment given the stock weights at every quarterly rebalancing date and assuming 10% of the average daily dollar traded volume can be traded. The nominal amount considered for the initial investment equals USD 1bn and is deflated back in time in line with the return of the cap-weighted market-index. The average 95th percentile of DTT is reported which is the time series average of the cross-sectional 95th percentile of DTT at each quarterly rebalancing. Data Source: CRSP, Compustat.

USA Long-Term 31-Dec-1972 to 31-Dec-2014	Number of Stocks in the Universe											
	250			500			1,000			3,000		
	5th %ile	Average	95th %ile	5th %ile	Average	95th %ile	5th %ile	Average	95th %ile	5th %ile	Average	95th %ile
Panel A: One-Way Turnover												
Cap-Weighted	7.03%	7.69%	8.45%	6.39%	7.28%	8.26%	6.21%	7.34%	8.46%	6.09%	7.69%	9.31%
Equal-Weighted	22.83%	23.48%	24.12%	21.67%	22.42%	23.15%	22.43%	23.33%	24.19%	26.17%	27.29%	28.51%
Fundamental-Weighted	20.78%	21.48%	22.31%	20.17%	21.35%	22.54%	20.16%	21.77%	23.20%	20.15%	22.75%	24.97%
Panel B: Transaction Costs												
Cap-Weighted	0.04%	0.04%	0.04%	0.03%	0.04%	0.05%	0.03%	0.04%	0.05%	0.04%	0.05%	0.07%
Equal-Weighted	0.12%	0.13%	0.14%	0.13%	0.14%	0.14%	0.16%	0.17%	0.18%	0.34%	0.38%	0.42%
Fundamental-Weighted	0.11%	0.11%	0.12%	0.11%	0.12%	0.13%	0.11%	0.13%	0.14%	0.13%	0.16%	0.19%
Panel C: Days to Trade (95th %ile)												
Cap-Weighted	1.38	2.06	3.09	1.56	2.39	3.74	2.44	3.79	5.82	6.04	9.99	14.96
Equal-Weighted	2.65	3.56	4.70	4.20	5.35	6.80	10.45	12.74	15.66	89.02	107.30	130.49
Fundamental-Weighted	1.63	2.48	3.79	2.02	2.93	4.44	3.25	4.80	7.03	9.27	15.13	22.17

3. Analysing Smart Beta Strategies

Table 6: Performance Measures of Generic Alternative Weighting Schemes (USA Long Term Track Records (LTTR) – Long Term – 42 Years)
The time period of analysis is 31-Dec-1972 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. The generic alternative weighting schemes assessed in addition to the traditional market cap-weighting are equal weighting and fundamental weighting. Fundamental Weighting weights stocks based on fundamental variables such as Book Value, Sales, Earnings and Cash Earnings. Generic alternative weighted schemes are tested across various size of the underlying universe. From the 3,000 largest stocks in USA, universes comprising of 250, 500, 1,000 and 3,000 largest stocks are chosen and from each universe 1,000 random samples of 100 stocks are selected and weighted according the generic weighting scheme chosen. For each combination of universe size and the weighting scheme chosen from the 1,000 portfolios formed the implementation costs that correspond to 5th percentile, average and 95th percentile are reported here. For relative calculations we use the Cap-weighted benchmark for a given stock selection without accounting for transaction costs (i.e. using its gross returns). The transaction cost estimates use the spread estimates according to the year of the rebalancing – Range-Based spread until 1993 and Closing Quoted Spread from 1993 onwards. Net returns are obtained after accounting for transaction costs at each quarterly rebalancing by multiplying the change in weight of each stock between the final weight before rebalancing and the optimal weights after rebalancing, including stock deletions and additions. Data Source: CRSP, Compustat.

USA Long-Term 31-Dec-1972 to 31-Dec-2014	Number of Stocks in the Universe											
	250			500			1,000			3,000		
	5th %ile	Average	95th %ile	5th %ile	Average	95th %ile	5th %ile	Average	95th %ile	5th %ile	Average	95th %ile
Panel A: Gross Annualised Returns												
Cap-Weighted	9.30%	9.76%	10.24%	9.43%	10.16%	10.86%	9.52%	10.44%	11.40%	9.30%	10.67%	11.97%
Equal-Weighted	10.28%	10.68%	11.10%	11.18%	11.74%	12.30%	11.68%	12.32%	12.96%	12.44%	13.31%	14.23%
Fundamental-Weighted	10.37%	10.87%	11.36%	10.66%	11.47%	12.25%	10.79%	11.83%	12.83%	10.48%	12.14%	13.58%
Panel B: Gross Relative Returns												
Cap-Weighted	-	-	-	-	-	-	-	-	-	-	-	-
Equal-Weighted	0.52%	0.93%	1.31%	0.95%	1.58%	2.19%	1.06%	1.88%	2.67%	1.25%	2.65%	3.97%
Fundamental-Weighted	0.85%	1.11%	1.39%	0.89%	1.31%	1.74%	0.75%	1.38%	1.99%	0.57%	1.47%	2.24%
Panel C: Gross Sharpe Ratio												
Cap-Weighted	0.22	0.25	0.28	0.23	0.28	0.32	0.23	0.29	0.35	0.21	0.30	0.38
Equal-Weighted	0.29	0.31	0.34	0.35	0.38	0.42	0.38	0.42	0.47	0.41	0.47	0.53
Fundamental-Weighted	0.28	0.31	0.34	0.30	0.35	0.40	0.29	0.37	0.44	0.26	0.38	0.48
Panel D: Gross Information Ratio												
Cap-Weighted	-	-	-	-	-	-	-	-	-	-	-	-
Equal-Weighted	0.16	0.28	0.39	0.22	0.36	0.51	0.19	0.34	0.50	0.14	0.31	0.47
Fundamental-Weighted	0.26	0.35	0.45	0.25	0.39	0.56	0.19	0.38	0.59	0.10	0.34	0.58
Panel E: Net Annualised Returns												
Cap-Weighted	9.26%	9.72%	10.20%	9.39%	10.12%	10.82%	9.47%	10.40%	11.36%	9.25%	10.62%	11.92%
Equal-Weighted	10.15%	10.55%	10.97%	11.04%	11.60%	12.17%	11.50%	12.15%	12.79%	12.05%	12.93%	13.86%
Fundamental-Weighted	10.26%	10.75%	11.25%	10.55%	11.35%	12.14%	10.65%	11.70%	12.72%	10.34%	11.97%	13.41%
Panel F: Net Relative Returns												
Cap-Weighted	-0.04%	-0.04%	-0.04%	-0.05%	-0.04%	-0.03%	-0.05%	-0.04%	-0.03%	-0.07%	-0.05%	-0.04%
Equal-Weighted	0.39%	0.80%	1.18%	0.82%	1.44%	2.06%	0.88%	1.70%	2.50%	0.86%	2.27%	3.60%
Fundamental-Weighted	0.74%	0.99%	1.28%	0.77%	1.19%	1.62%	0.62%	1.25%	1.85%	0.40%	1.31%	2.07%

3. Analysing Smart Beta Strategies

Panel G: Net Sharpe Ratio												
Cap-Weighted	0.22	0.25	0.28	0.23	0.27	0.32	0.23	0.29	0.35	0.21	0.29	0.38
Equal-Weighted	0.28	0.30	0.33	0.34	0.37	0.41	0.37	0.41	0.46	0.39	0.45	0.50
Fundamental-Weighted	0.27	0.30	0.34	0.29	0.34	0.39	0.29	0.36	0.43	0.25	0.37	0.47
Panel H: Net Information Ratio												
Cap-Weighted	-1.41	-1.28	-1.12	-1.42	-1.21	-0.97	-1.47	-1.23	-0.96	-1.52	-1.25	-0.91
Equal-Weighted	0.12	0.24	0.36	0.19	0.33	0.48	0.16	0.31	0.47	0.10	0.27	0.43
Fundamental-Weighted	0.22	0.31	0.41	0.22	0.36	0.52	0.16	0.35	0.55	0.07	0.30	0.54

The results suggest that net performance is very close to gross performance in the case of the large-cap universes (250 stocks or 500 stocks). For the broad market universe (3,000 stocks), the wedge between gross and net performance becomes more pronounced, which leads to considerably higher cost levels.

Table 7 reports several liquidity metrics for our generic strategies.

Table 7: Liquidity Measures of Generic Alternative Weighting Schemes (USA Long Term Track Records (LTTR) – Long Term – 42 Years) The time period of analysis is 31-Dec-1972 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. The generic alternative weighting schemes assessed in addition to the traditional market cap-weighting are equal weighting and fundamental weighting. Fundamental Weighting weights stocks based on fundamental variables such as Book Value, Sales, Earnings and Cash Earnings. Generic alternative weighted schemes are tested across various size of the underlying universe. From the 3,000 largest stocks in USA, universes comprising of 250, 500, 1,000 and 3,000 largest stocks are chosen and from each universe 1,000 random samples of 100 stocks are selected and weighted according the generic weighting scheme chosen. For each combination of universe size and the weighting scheme chosen from the 1,000 portfolios formed the implementation costs that correspond to 5th percentile, average and 95th percentile are reported here. The transaction cost estimates use the spread estimates according to the year of the rebalancing – Range-Based spread until 1993 and Closing Quoted Spread from 1993 onwards. Weighted average market capitalisation (in USD millions) is obtained for each quarter by summing the product of the weight of the stock at each quarterly rebalancing by its corresponding total market capitalisation (in USD millions) and then a time-series average across quarters is calculated and reported. The Amihud measure for a given stock is calculated each day by dividing the absolute value of the stock's return by the traded volume. The Amihud measure is multiplied by 10^9 for ease of presentation. The quarterly average is then obtained by averaging the daily values for which the Amihud measure exists. The weighted average Amihud measure is then obtained for each quarter and a time series average across all quarters are calculated and reported. Weighted average spread is obtained for each quarter by summing the product of the weight of the stock at each quarterly rebalancing by its corresponding spread estimate (Range-Based spread until 1993 and Closing Quoted spread from 1993 onwards) and then a time-series average across quarters is calculated and reported. Extreme liquidity measures are obtained by sorting the universe by the corresponding liquidity measure and report the measure corresponding to the stock that represents the 95% cumulative weight in the corresponding smart beta strategy. Data Source: CRSP, Compustat.

USA Long-Term 31-Dec-1972 to 31-Dec-2014	Number of Stocks in the Universe											
	250			500			1,000			3,000		
	5th %ile	Average	95th %ile	5th %ile	Average	95th %ile	5th %ile	Average	95th %ile	5th %ile	Average	95th %ile
Panel A: Weighted Average Market Capitalisation												
Cap-Weighted	59,660	65,237	69,735	42,876	56,620	68,213	31,748	49,561	68,476	17,127	40,604	70,231
Equal-Weighted	27,756	29,524	31,052	16,471	19,102	21,563	9,518	12,047	14,664	3,509	5,307	7,554
Fundamental-Weighted	53,659	58,582	62,035	36,800	49,427	60,201	27,209	42,826	59,981	14,907	34,857	65,283
Panel B: Weighted Average Spread												
Cap-Weighted	0.37%	0.38%	0.39%	0.38%	0.40%	0.42%	0.40%	0.42%	0.45%	0.43%	0.49%	0.55%
Equal-Weighted	0.42%	0.43%	0.43%	0.45%	0.46%	0.47%	0.52%	0.54%	0.57%	0.90%	0.98%	1.06%
Fundamental-Weighted	0.39%	0.40%	0.41%	0.41%	0.42%	0.44%	0.43%	0.45%	0.48%	0.47%	0.53%	0.59%

3. Analysing Smart Beta Strategies

Panel C: Weighted Average Amihud Measure												
Cap-Weighted	3.00	3.72	4.51	4.52	6.41	8.79	8.11	12.95	19.71	19.04	39.35	66.88
Equal-Weighted	5.84	7.40	8.95	11.26	14.65	18.56	29.16	41.98	62.91	261.24	358.51	523.56
Fundamental-Weighted	3.54	4.55	5.69	5.42	7.72	10.39	9.84	15.57	23.31	32.04	62.55	104.14
Panel D: Extreme Low (95% Cumulative Weight) Market Capitalisation												
Cap-Weighted	8,943	9,310	9,722	4,965	5,394	5,863	2,490	3,013	3,548	864	1,285	1,851
Equal-Weighted	6,708	6,836	6,992	3,162	3,321	3,493	1,216	1,286	1,367	144	163	184
Fundamental-Weighted	8,133	8,425	8,740	4,522	4,882	5,264	2,136	2,556	3,032	665	960	1,350
Panel E: Extreme High (95% Cumulative Weight) Spread												
Cap-Weighted	0.70%	0.72%	0.74%	0.74%	0.77%	0.80%	0.78%	0.85%	0.92%	0.89%	1.09%	1.33%
Equal-Weighted	0.78%	0.80%	0.82%	0.85%	0.88%	0.91%	1.09%	1.19%	1.29%	2.61%	2.89%	3.18%
Fundamental-Weighted	0.74%	0.76%	0.77%	0.78%	0.81%	0.84%	0.84%	0.90%	0.97%	1.00%	1.22%	1.48%
Panel F: Extreme High (95% Cumulative Weight) Amihud Measure												
Cap-Weighted	8.59	11.20	14.91	15.55	22.95	32.27	29.47	47.47	70.31	53.83	135.49	250.41
Equal-Weighted	16.34	24.32	33.53	33.38	48.21	71.91	97.66	137.02	192.53	1040.04	1397.41	1854.19
Fundamental-Weighted	9.77	16.04	25.82	17.98	27.51	37.51	37.64	57.91	84.52	103.27	229.99	373.97

The results in Table 7 suggest again that implementation issues remain well-behaved in the narrow large-cap universes but become more pronounced in the broader universe.

Overall, across our analysis of generic strategies drawing on different types of stock universes, we find pronounced dependence of conclusions about transaction cost levels and implementation challenges on the stock universe used for implementation. While it is common to see broad brush statements about investability hurdles of particular smart beta strategies, our results provide clear evidence that conclusions heavily depend on the universe under consideration. It is obviously sensible to consider that implementing a strategy such as equal weighting in a universe of mega cap or large-cap stocks will be very different in terms of trading costs than implementing such a strategy in a broad market universe that includes small cap stocks. Our results on generic strategies show indeed that cost and investability metrics differ tremendously across universes. For example, the results in Table 5 suggest that for portfolios built from the top 250 stocks by market cap, we obtain days to trade measures of 3.56 days compared to 2.06 for the cap-weighted portfolios in the same universe. Moreover, the estimate of average annualised transaction costs is 0.13% for the equal-weighted portfolios compared to 0.04% for the cap-weighted portfolios in the 250-stock universe. When looking at portfolios formed from the broad universe (the top 3,000 stocks by market cap), we get strikingly different results. The days to trade measure reaches more than 100 for equal-weighted portfolios compared to about 10 for cap-weighted portfolios. Estimated transaction costs are 0.38% for

3. Analysing Smart Beta Strategies

equal-weighted portfolios compared to 0.05% for cap-weighted portfolios. Thus, an equal weighting strategy indeed looks extremely challenging to implement for the broad universe, but implementation measures are rather well-behaved for the large-cap universe. Given such differences, it makes little sense to make statements about the investability of equal-weighing per se without considering the universe it is implemented for.

3.2 Scientific Beta Index Strategies

An analysis of generic strategies allows us to vary the universe and also generate multiple random portfolios which allows for insights into the distribution of outcomes. However, from a practical perspective, the key question is what the replication costs and implementation hurdles are for the indices that are constructed with typical universes used by strategy providers. In this subsection, we analyse commercially available index strategies reflecting three types of popular smart beta strategies, namely a Minimum Volatility strategy, an (adjusted) Equal-Weighting Strategy and a Multi-Factor index strategy. We use long term track records corresponding to indices from Scientific Beta.[9] The corresponding data is available over long time periods, including the quarterly holdings data which allows us to estimate stock level transaction costs and apply them to the strategies. The USA Long Term Track Records of Scientific Beta allow us to conduct an analysis over the long term (42 years).

We analyse three smart beta strategies, namely the Scientific Beta Efficient Minimum Volatility Index, the Maximum Deconcentration Index and the Multi-Beta Multi-Strategy Index for USA Long Term Track Records in the presence of transaction costs. These three strategies are chosen as they are the most popular among investors. The Efficient Minimum Volatility and Maximum Deconcentration strategies are diversification strategies that aim to provide a better risk-adjusted return compared to cap-weighted benchmark whereas the Multi-Beta Multi-Strategy Index is a multi-factor index that provides exposure to some of the well-known risk factors thus outperforming the cap-weighted benchmark by capturing the associated risk premia. The SciBeta USA universe comprises of 500 largest stocks in USA. The long term data are available for the entire period of our analysis (i.e. 42 years). The following are the brief description of each of these smart beta strategies.

Scientific Beta Efficient Minimum Volatility:

The Minimum Volatility strategy is a well-accepted solution among investors seeking low risk equity investments. Also, like many other alternative weighting schemes, the Minimum Volatility strategy gained popularity as an equity indexing scheme following increasing recognition of the shortcomings of cap-weighted indices, and notably their failure to represent a good proxy for risk/reward efficient portfolios.[10] Substantial empirical evidence exists to show that Minimum Volatility portfolios outperform the cap-weighted benchmark with lower risk in the US equity universe (Chan et al., 1999; Schwartz, 2000; Jagannathan and Ma, 2003; Clarke, de Silva and Thorley, 2006; DeMiguel et al., 2009a). Similar results are confirmed by Geiger and Plagge (2007), Nielsen and Aylursubramanian (2008), and Poullaouec (2008) for Minimum Volatility portfolios in global equity markets. The true minimum volatility portfolio lies on the efficient frontier and coincides with

9 - Scientific Beta is a smart beta index provider that had been set up by EDHEC-Risk Institute as part of its policy of transferring know-how to the industry. ERI Scientific Beta is an initiative which aims to favour the adoption of the latest advances in smart beta design and implementation by the investment industry. Data for the indices we use in this paper is available through the website www.scientificbeta.com or through data providers such as Bloomberg.

10 - Cap-weighting leads to the optimal portfolio (in the sense of the Maximum Sharpe Ratio portfolio) according to Sharpe's (1964) Capital Asset Pricing Model (CAPM). However, critics (for example, Roll (1977), Fama and French (2004), among others) argue that most of the CAPM's assumptions – such as all investors having identical expectations and time horizons, and all securities being tradable – do not hold.

3. Analysing Smart Beta Strategies

the optimal portfolio of Modern Portfolio Theory (the tangency portfolio) if, and only if, expected returns are identical across all stocks. However, due to the presence of estimation risk affecting the input parameters, the minimum volatility portfolio is an attractive strategy because there is no need to estimate expected returns (only risk parameters need to be estimated). Thus, minimum volatility strategies can, in practice, hope to be decent proxies of truly efficient portfolios. Scientific Beta Efficient Minimum Volatility weighting scheme aims to provide an implementable proxy for the Minimum Volatility portfolio postulated by Modern Portfolio Theory (MPT). It uses state-of-the-art risk parameter estimates and provides an improvement upon standard Minimum Volatility portfolios by introducing constraints on portfolio concentration. Detailed discussion of the benefits and implementation of the Efficient Minimum Volatility strategy can be found in Goltz and Lodh (2015).[11]

Scientific Beta Maximum Deconcentration:

The Maximum Deconcentration strategy originates from equally-weighted portfolios in which the dollar proportion of each asset is set to be identical. Platen and Rendek (2010), Dash and Loggie (2008) and DeMiguel et al. (2009b) among others have found that equally-weighted portfolios deliver improved Sharpe ratios compared to cap-weighted portfolios. A plain vanilla equal weighting strategy, however, usually has high turnover and liquidity problems. Maximum Deconcentration portfolios aim at getting as close as possible to equal weights while respecting practical investment constraints, notably implementation constraints (turnover control and liquidity rules) and, optionally, additional constraints on risk exposures. Maximum Deconcentration can be perceived as a generalisation of a simple equal weighting scheme: the aim being to maximise the effective number of stocks. Detailed discussion of the benefits and implementation of the maximum deconcentration strategy can be found in Gonzalez et al. (2015).[12]

Scientific Beta Multi-Beta Multi-Strategy (4-Factor) EW:

Multi-factor indices represent the recent innovation in the field of factor investing and the Scientific Beta Multi-Beta Multi-Strategy (4-Factor) EW index is one of Scientific Beta's flagship multi-factor indices that aim to provide balanced exposure to the rewarded risk factors in a well diversified way. It is an equal-weighted combination of the four smart factor indices represented by Diversified Multi-Strategy factor indices that provide access to the well known factors – Value, Momentum, Size and Low Volatility (Scientific Beta Value Multi-Strategy Indices, Scientific Beta Momentum Multi-Strategy Indices, Scientific Beta Mid Cap Multi-Strategy Indices, and Scientific Beta Low Vol Multi-Strategy Indices). It uses Scientific Beta's well diversified smart factor indices as building blocks. Smart factor indices follow Scientific Beta's original 'Smart Beta 2.0' approach to diversification, which separates the stock selection process and applies a diversification-based weighting scheme to ensure the obtention of factor exposure without losing diversification benefits. The diversified multi-strategy weighting scheme is an equal-weighted combination of the following five weighting schemes – Maximum Deconcentration, Diversified Risk Weighted, Maximum Decorrelation, Efficient Minimum Volatility and Efficient Maximum Sharpe Ratio. Detailed discussion of the benefits and

11 - Goltz and Lodh (2015) can be accessed online at http://www.scientificbeta.com/download/file/scientific-beta-efficient-min-volatility-indices.

12 - Gonzalez et al. (2015) can be accessed online at http://www.scientificbeta.com/download/file/scientific-beta-max-deconcentration-indices.

3. Analysing Smart Beta Strategies

implementation of the Multi-Beta Multi-Strategy (4-Factor) EW index can be found in Amenc et al (2016).[13]

Table 8 summarises the results of a transaction cost analysis for USA Long Term track records over long term (42 years).

Table 8: Transaction Costs, Performance and Liquidity Measures of Popular Smart Beta Strategies (USA Long Term Track Records (LTTR) – Long Term – 42 Years)
The time period of analysis is 31-Dec-1972 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. The strategies considered for this analysis are SciBeta USA LTTR Efficient Minimum Volatility Index, SciBeta USA LTTR Maximum Deconcentration Index and SciBeta USA LTTR Multi-Beta Multi-Strategy (4-Factor) EW Index. The rebalancing frequency of the strategies is quarterly subject to buffer, turnover control and liquidity capping rules. One-way Turnover is half the sum of absolute deviations of individual weights between the end of a quarter and the beginning of the following quarter. One-way Turnover is annualised for each quarter by multiplying it by 4. The transaction cost estimates use the spread estimates according to the year of the rebalancing – Range-Based spread until 1993 and Closing Quoted Spread from 1993 onwards. The reported transaction cost estimates are the difference between the annualised gross and net returns. Net returns are obtained after accounting for transaction costs at each quarterly rebalancing by multiplying the change in weight of each stock (including stock deletions and additions) between the final weight before rebalancing and the optimal weights after rebalancing by a half spread. Days to trade (DTT) refers to the time it takes to trade an initial investment given the stock weights at every quarterly rebalancing date, assuming 10% of the average daily dollar traded volume can be traded. The nominal amount considered for the initial investment equals USD 1bn and is deflated back in time in line with the return of the cap-weighted market-index. The average 95th percentile of DTT is reported which is the time series average of the cross-sectional 95th percentile of DTT at each quarterly rebalancing. For relative calculations we use the Cap-weighted benchmark for a given stock selection without accounting for transaction costs (i.e. using its gross returns). Weighted average market capitalisation (in USD millions) is obtained for each quarter by summing the product of the weight of the stock at each quarterly rebalancing by its corresponding total market capitalisation (in USD millions) and then a time-series average across quarters is calculated and reported. The Amihud measure for a given stock is calculated each day by dividing the absolute value of the stock´s return by the traded volume. The Amihud measure is multiplied by 10^9 for ease of presentation. The quarterly average is then obtained by averaging the daily values for which the Amihud measure exists. The weighted average Amihud measure is then obtained for each quarter and a time series average across all quarters are calculated and reported. Weighted average spread is obtained for each quarter by summing the product of the weight of the stock at each quarterly rebalancing by its corresponding spread estimate (Range-Based spread until 1993 and Closing Quoted spread from 1993 onwards) and then a time-series average across quarters is calculated and reported. Extreme liquidity measures are obtained by sorting the universe by the corresponding liquidity measure and report the measure corresponding to the stock that represents the 95% cumulative weight in the corresponding smart beta strategy. Data Source: CRSP, Scientific Beta.

	SciBeta Indices							
USA LTTR Long-Term 31-Dec-1972 to 31-Dec-2014	Broad Cap-Weighted		Efficient Minimum Volatility		Maximum Deconcentration		Multi-Beta Multi-Strategy 4-Factor EW	
Panel A – Implementation Costs								
One-Way Turnover	2.63%		30.02%		20.29%		28.89%	
Transaction Costs	0.01%		0.18%		0.13%		0.17%	
Days to Trade (95 %ile)	0.39		2.19		2.15		3.50	
Panel B – Performance Measures								
	Gross	Net	Gross	Net	Gross	Net	Gross	Net
Annualised Returns	10.32%	10.30%	12.80%	12.63%	12.70%	12.57%	14.25%	14.08%
Relative Returns	-	-0.01%	2.48%	2.31%	2.38%	2.26%	3.93%	3.77%
Sharpe Ratio	0.29	0.29	0.52	0.51	0.43	0.43	0.58	0.57
Information Ratio	-	-1.32	0.48	0.44	0.56	0.53	0.77	0.73
Panel C – Liquidity Measures								
	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)
Market Capitalisation	48,013	3,209	14,053	1,430	11,746	1,267	10,314	1,232
Spread	0.41%	0.81%	0.47%	0.95%	0.53%	1.15%	0.50%	1.06%
Amihud Measure	7.80	60.80	28.74	218.20	41.43	350.91	51.41	979.44

13 - Amenc et al. (2016) can be accessed online at http://www.scientificbeta.com/download/file/scientific-beta-multi-beta-multi-strategy-indices.

3. Analysing Smart Beta Strategies

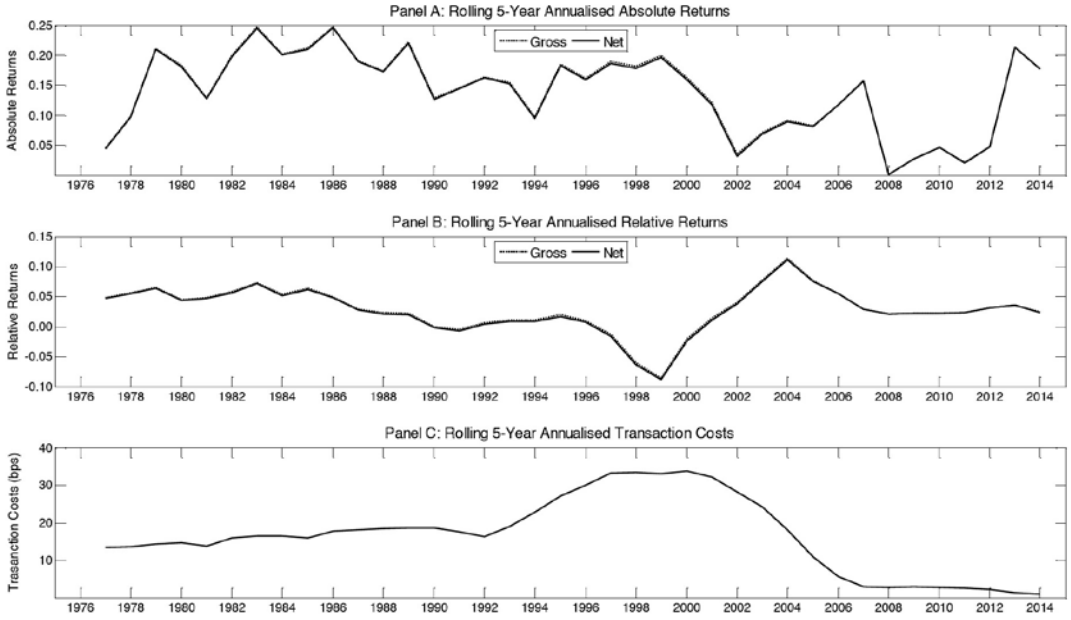
It appears that transaction costs have a relatively mild impact on performance. It should be noted that the time period covers data going back as far as the 1970s and thus much of the sample does not yet capture the recent decline in transaction costs. It may be interesting to see whether the difference between gross returns and net returns was more noticeable in the earlier time periods compared to the more recent time periods.

Exhibit 3 provides some perspective on this question. It shows five-year rolling window returns, both net and gross returns. For reasons of readability, the graphs show the average returns across the three strategies analysed in Table 8.

It is rather clear from inspecting the lines for net and gross returns that transaction costs hardly alter the returns of these strategies. In particular, while one may expect that that the difference is somewhat more noticeable in the earlier years of the sample, net returns are indeed hardly visually distinguishable from gross returns across the entire sample period.

Apart from the significant increase of market liquidity during the past years, the decline in transaction costs can be also attributed to market microstructure changes such as the reduction of the minimum tick size. Table 9 reports performance measures along with transaction costs from 31 December 1972 to 31 December 1996 when the minimum

Exhibit 3: Rolling Window Analysis of Popular Smart Beta Strategies (Average across three Strategies; USA Long Term Track Records) The exhibit presents the average annualised gross and net returns, gross and net relative returns and transaction costs of the three smart beta strategies – the SciBeta USA LTTR Efficient Minimum Volatility Index, the SciBeta USA LTTR Maximum Deconcentration Index and the SciBeta USA LTTR Multi-Beta Multi-Strategy (4-Factor) EW Index using a rolling 5-year window with 1-year step size. Panel A presents the gross and net absolute returns; Panel B presents the gross and net relative returns; Panel C presents the transaction costs. The average returns/costs of the three smart beta indices each year are plotted. The time period of analysis is 31-Dec-1972 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. The transaction costs estimates use the spread estimates according to the year of the rebalancing – Range-Based spread until 1993 and Closing Quoted Spread from 1993 onwards. The reported transaction cost estimates are the difference between the annualised gross and net returns. Net returns are obtained after accounting for transaction costs at each quarterly rebalancing by multiplying the change in weight of each stock between the final weight before rebalancing and the optimal weights after rebalancing, including stock deletions and additions. Data Source: CRSP, Scientific Beta.



3. Analysing Smart Beta Strategies

Table 9: Performance and Transaction costs of Popular Smart Beta Strategies under various market microstructure regimes (USA Long Term Track Records (LTTR) – Long Term – 42 Years)
The time period of analysis is 31-Dec-1972 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. The strategies considered for this analysis are the SciBeta USA LTTR Efficient Minimum Volatility Index, the SciBeta USA LTTR Maximum Deconcentration Index and the SciBeta USA LTTR Multi-Beta Multi-Strategy (4-Factor) EW Index. The transaction cost estimates use the spread estimates according to the year of the rebalancing – Range-Based spread until 1993 and Closing Quoted Spread from 1993 onwards. The reported transaction cost estimates are the difference between the annualised gross and net returns. Net returns are obtained after accounting for transaction costs at each quarterly rebalancing by multiplying the change in weight of each stock (including stock deletions and additions) between the final weight before rebalancing and the optimal weights after rebalancing by half spread. In the US stock market, the minimum tick size is reduced form 1/8th to 1/16th in 1997, and from 1/16th to 1/100th (or decimalisation) in 2001. Data Source: CRSP, Scientific Beta.

Panel A - Period in which minimum tick size is 1/8th

	SciBeta Indices							
USA LTTR Long-Term 31-Dec-1972 to 31-Dec-1996	Broad Cap-weighted		Efficient Minimum Volatility		Maximum Deconcentration		Multi-Beta Multi-Strategy 4-Factor EW	
	Gross	Net	Gross	Net	Gross	Net	Gross	Net
Annualised Returns	12.21%	12.20%	14.78%	14.56%	14.40%	14.25%	16.36%	16.15%
Transaction Costs	0.02%		0.22%		0.14%		0.20%	
Relative Returns	-	-0.02%	2.56%	2.34%	2.18%	2.04%	4.14%	3.94%
Sharpe Ratio	0.32	0.32	0.58	0.57	0.49	0.48	0.68	0.66
Information Ratio	-	-1.47	0.57	0.52	0.57	0.53	0.88	0.84

Panel B - Period in which minimum tick size is 1/100th

	SciBeta Indices							
USA LTTR Long-Term 31-Dec-2001 to 31-Dec-2014	Broad Cap-weighted		Efficient Minimum Volatility		Maximum Deconcentration		Multi-Beta Multi-Strategy 4-Factor EW	
	Gross	Net	Gross	Net	Gross	Net	Gross	Net
Annualised Returns	6.76%	6.75%	10.15%	10.11%	9.65%	9.62%	10.80%	10.76%
Transaction Costs	< 0.01%		0.03%		0.03%		0.04%	
Relative Returns	-	0.00%	3.39%	3.36%	2.89%	2.86%	4.04%	4.01%
Sharpe Ratio	0.27	0.27	0.5	0.49	0.38	0.38	0.49	0.49
Information Ratio	-	-0.95	0.83	0.82	0.76	0.75	1.09	1.08

tick size was 1/8th of a dollar in Panel A, and from 31 December 2001 to 31 December 2014 when the minimum tick size was 1 cent in Panel B. It is clear that the transaction costs are significantly reduced on average in the post decimalisation period compared to the period when the minimum tick size was 1/8th.

3.3 Impact of Implementation Rules

Typically, strategy providers apply certain implementation rules or constraints to the index construction process in order to provide a more practical index outcome by reducing turnover and by improving liquidity, thereby reducing the transaction costs. Implementation rules normally take the form of liquidity or turnover constraints. A typical liquidity constraint would be to control the size of rebalancing or total index weight invested in one stock relative to the stock size, the reason being to prevent relatively large trades in smaller, illiquid stocks that could be difficult to execute at a reasonable cost. Typical turnover constraints are to directly impose limits on turnover, limit the frequency of rebalancing or rebalancing only above some threshold amount of turnover. The latter has been shown by Leland (1999) to substantially reduce

3. Analysing Smart Beta Strategies

turnover in the presence of transaction costs compared to an unconstrained rebalancing strategy. The following two sections explain the two types of implementation rules that are included in the indices provided by Scientific Beta that were analysed above (i.e. turnover rules and liquidity rules). These sections are followed by an empirical assessment of the impact these implementation rules have on transaction costs, by separating the index construction process to before and after the application of the implementation rules.

3.3.1 Turnover Constraints

Systematically rebalancing a portfolio could lead to high turnover. In order to avoid such an undesirable effect, we employ a threshold-based rebalancing method which is conditional in nature. At the rebalancing date, the portfolio is assigned optimal weights only if the impact of the potential weight changes is substantial, showing that a significant amount of new information has arrived since the last rebalancing.

At each review, we compute the full potential set of new weights w_{POT} as per the weighting scheme of the corresponding smart beta strategy. We then compare them to the current weights w_{CURR} at the close of business day prior to rebalancing. The potential two-way turnover δ is then computed as follows:

$$\delta = \sum_{i=1}^{N} \left| w_{CURR}^{i} - w_{POT}^{i} \right| \geq \delta_{THRES}.$$

If the value obtained for δ is greater than a calibrated threshold value δ_{THRES}, the optimal weights are applied, or else only the index constitution is updated, by integrating incoming stocks and removing outgoing stocks (which result from quarterly universe reconstitution).

The threshold level δ_{THRES} is calibrated on back-tests for the first 10 years of portfolio back tested history in the following manner: we optimise and calculate the index for all δ spanning from 0% to 100% (by 5% increments). We pick the smallest δ that has led to a resulting one-way annual turnover under 30% in that back test period. If no such threshold can be found, then the turnover target is loosened to 35%, then 40%, and so on (by 5% increments). Irrespective of whether or not the threshold mentioned above is reached, new weights are used if the index has not been rebalanced optimally for seven consecutive quarters. This ensures that optimal weights are imposed at least every two years. The intent of this rule is to avoid rebalancing when deviations of new optimal weights from the current weights are relatively small.

3.3.2 Liquidity Constraints

In order to improve the liquidity of the portfolio, additional adjustments of weights are implemented to achieve two objectives – one is to limit liquidity issues that may arise upon investing in the portfolio and another is to limit the liquidity issues that may occur upon rebalancing the portfolio. The principle used to make such adjustments is to impose an upper bound for the weight of a stock and for the weight change at rebalancing, relative to the market cap weight of the stock in its universe.

Specifically, liquidity adjustments consist of two rules:

- Holding Capacity Rule – The stock's weight is capped at a maximum of 10 times the free-float adjusted market cap weight to avoid big investment in the smallest stocks.
- Trading Capacity Rule – The change in weight for each stock is capped at its

3. Analysing Smart Beta Strategies

free-float adjusted market cap weight to avoid large rebalancing in small stocks.

- After capping weights and weight changes following the two adjustments above, index weights are renormalised so that they sum again to one.

3.3.3 Empirical assessment of the impact of implementation rules on Transaction Costs

In this section, we assess the impact of the implementation rules on the transaction costs of the popular smart beta strategies discussed in the previous sections – the Efficient Minimum Volatility index and the Maximum Deconcentration index. The Multi-Beta Multi-Strategy 4-Factor EW index has many component indices and the implementation rules are applied at the component index level so, for the sake of brevity, the index is omitted in the analysis here. The index construction process is split and indices are constructed at each of the following intermediary steps.

1. Optimal weights determined by the appropriate weighting scheme as it is used to produce the index (no rules applied);

2. Turnover rules alone are applied and the resulting weights are used to construct the index;

3. Liquidity rules are also applied and the resulting weights are used to construct the index.

Two new measures are added in this analysis corresponding to the two liquidity rules applied – holding constraints and trading constraints. The holding multiple of each stock in a quarter is the ratio of each stock's weight in the index to its weight in the cap-weighted index. The average of the 95th percentile value in each quarter is reported. The trading multiple of each stock in a quarter is the ratio of weight change of each stock in the index to its weight in the cap-weighted index. The average of the 95th percentile value in each quarter is reported.

Table 10 shows the results of the assessment of transaction costs, performance measures and liquidity measures of the two indices at each intermediary index construction process over the long term for the USA LTTR universe. It is clear that the turnover rule substantially reduces the turnover in the case of the Efficient Minimum Volatility index without losing much performance. In the case of the Maximum Deconcentration index, the turnover rule does not bind as the annualised turnover is less than 30%, even without any constraints. The liquidity rules improve the liquidity of both indices. Overall, after the application of the implementation rules, we witness a significant improvement in liquidity coupled with a reduction in turnover thus leading to a reduction in transaction costs of both smart beta strategies.

3. Analysing Smart Beta Strategies

Table 10: Impact of Turnover and Liquidity Rules on Scientific Beta Strategies – USA LTTR Long Term
The time period of analysis is 31-Dec-1972 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. The strategies considered for this analysis are the SciBeta USA LTTR Efficient Minimum Volatility Index and the SciBeta USA LTTR Maximum Deconcentration Index, constructed with and without implementation rules. The rebalancing frequency of the strategies is quarterly subject to buffer, turnover control and liquidity capping rules. One-way Turnover is half the sum of absolute deviations of individual weights between the end of a quarter and the beginning of the following quarter. One-way Turnover is annualised for each quarter by multiplying it by 4. The transaction cost estimates use the spread estimates according to the year of the rebalancing – Range-Based spread until 1993 and Closing Quoted Spread from 1993 onwards. The reported transaction cost estimates are the difference between the annualised gross and net returns. Net returns are obtained after accounting for transaction costs at each quarterly rebalancing by multiplying the change in weight of each stock between the final weight before rebalancing and the optimal weights after rebalancing, including stock deletions and additions. Days to trade (DTT) refers to the time it takes to trade an initial investment given the stock weights at every quarterly rebalancing date, assuming 10% of the average daily dollar traded volume can be traded. The nominal amount considered for the initial investment equals USD 1bn and is deflated back in time in line with the return of the cap-weighted market-index. The average 95th percentile of DTT is reported which is the time series average of the cross-sectional 95th percentile of DTT at each quarterly rebalancing. For relative calculations we use the Cap-weighted benchmark for a given stock selection without accounting for transaction costs (i.e. using its gross returns). Weighted average market capitalisation (in USD millions) is obtained for each quarter by summing the product of the weight of the stock at each quarterly rebalancing by its corresponding total market capitalisation (in USD millions) and then a time-series average across quarters is calculated and reported. The Amihud measure for a given stock is calculated each day by dividing the absolute value of the stock´s return by the traded volume. The Amihud measure is multiplied by 10^9 for ease of presentation. The quarterly average is then obtained by averaging the daily values for which the Amihud measure exists. The weighted average Amihud measure is then obtained for each quarter and a time series average across all quarters are calculated and reported. Weighted average spread is obtained for each quarter by summing the product of the weight of the stock at each quarterly rebalancing by its corresponding spread estimate (Range-Based spread until 1993 and Closing Quoted spread from 1993 onwards) and then a time-series average across quarters is calculated and reported. Extreme liquidity measures are obtained by sorting the universe by the corresponding liquidity measure and the measure corresponding to the stock that represents the 95% cumulative weight in the corresponding smart beta strategy is reported. Holding Multiple of each stock in a quarter is the ratio of the weight of each stock in the index to the weight of the stock in the cap-weighted index. The average of the 95th percentile value in each quarter is reported. Trading Multiple of each stock in a quarter is the ratio of the weight change of each stock in the index to the weight of the stock in the cap-weighted index. The average of the 95th percentile value in each quarter is reported. Data Source: CRSP, Scientific Beta.

	SciBeta USA LTTR Indices													
USA LTTR Long-Term 31-Dec-1972 to 31-Dec-2014	Broad Cap-Weighted		Efficient Minimum Volatility						Maximum Deconcentration					
			Before Turnover and Liquidity Rules		After Turnover but Before Liquidity Rules		After Turnover and Liquidity Rules		Before Turnover and Liquidity Rules		After Turnover but Before Liquidity Rules		After Turnover and Liquidity Rules	
Panel A - Implementation Costs														
One-Way Turnover	2.63%		54.57%		37.96%		30.02%		23.22%		23.22%		20.29%	
Transaction Costs	0.01%		0.38%		0.29%		0.18%		0.15%		0.15%		0.13%	
Days to Trade (95th %ile)	0.39		3.14		3.13		2.19		3.47		3.47		2.15	
Holding Multiple (99th %ile)	1.00		37.29		36.85		11.04		38.94		38.94		10.79	
Trading Multiple (99th %ile)	0.01		15.53		9.64		1.30		7.92		7.92		1.16	
Panel B – Performance Measures														
	Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
Annualised Returns	10.32%	10.30%	13.09%	12.71%	12.98%	12.69%	12.80%	12.63%	12.90%	12.75%	12.90%	12.75%	12.70%	12.57%
Relative Returns	-	-0.01%	2.77%	2.40%	2.66%	2.37%	2.48%	2.31%	2.59%	2.43%	2.59%	2.43%	2.38%	2.26%
Sharpe Ratio	0.29	0.29	0.55	0.52	0.54	0.52	0.52	0.51	0.44	0.43	0.44	0.43	0.43	0.43
Information Ratio	-	-1.32	0.50	0.43	0.48	0.42	0.48	0.44	0.57	0.53	0.57	0.53	0.56	0.53
Panel C– Liquidity Measures														
	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)
Market Capitalisation	48,013	3,209	13,643	1,348	13,684	1,350	14,053	1,430	11,527	1,116	11,527	1,116	11,746	1,267
Spread	0.41%	0.81%	0.48%	1.00%	0.48%	1.00%	0.47%	0.95%	0.56%	1.26%	0.56%	1.26%	0.53%	1.15%
Amihud Measure	7.80	60.80	46.43	307.69	45.15	456.28	28.74	218.20	74.86	705.31	74.86	705.31	41.43	350.91

3. Analysing Smart Beta Strategies

3.4 Transaction Costs Assessment of International Indices

It is interesting to extend our analysis of transaction costs to the smart beta indices in other non-US Developed countries. In this section, we perform an analysis similar to that in Section 3.2 for the same three Smart Beta Strategies, namely Efficient Minimum Volatility, Maximum Deconcentration and Multi-Beta Multi-Strategy 4-Factor EW in the Scientific Beta Developed-ex USA universe. We report the transaction costs, performance measures and liquidity measures of these indices. As long-term data is not available for non-US universes, we restrict our analysis to a 10-year history. In order to compare the results of the Developed ex USA and the USA universe, we also report similar performance measures for the same 10-year history for US indices. The SciBeta USA universe comprises the 500 largest stocks in the USA, and the Developed ex USA universe roughly comprises the 1,500 largest stocks in the universe. The results are reported in Table 11 and Table 12 for Developed ex USA and USA respectively. Given the difference in liquidity and the depth of the US and other Developed markets, it is unsurprising to see that the transaction costs of the indices in the Developed ex USA universe are a magnitude larger than those of the USA universe. A graphical comparison between spread levels for the USA and developed ex USA universe across different size groups can be found in the appendix (available upon request).

However, despite the higher level of transaction costs in developed markets, cost levels are not high enough to erode the relative returns of the smart beta indices. In particular, cost levels for the three strategies analysed here remain at annualised levels of less than 20 bps. Compared to the annualised gross returns that often exceed levels of 2%, these costs are not sufficient to significantly alter the conclusions on the performance of such strategies. However, it should be stressed that these cost levels apply to investable indices which are constructed from liquid stock universes (i.e. large- and mid-cap stocks) and come with built-in implementation constraints. We have seen in the analysis of US data that the use of a liquid universe and implementation rules help maintain costs and liquidity issues at manageable levels. Given the higher cost levels in international markets, considering issues such as using an appropriate universe and applying implementation rules becomes even more relevant when designing strategies in these markets, compared to designing strategies for the US market.

3. Analysing Smart Beta Strategies

Table 11: Transaction Costs, Performance and Liquidity Measures of Popular Smart Beta Strategies (Developed World ex United States – Short Term – 10 Years)
The time period of analysis is 31-Dec-2004 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. The strategies considered for this analysis are the SciBeta Developed ex USA Efficient Minimum Volatility Index, the SciBeta Developed ex USA Maximum Deconcentration Index and the SciBeta Developed ex USA Multi-Beta Multi-Strategy (4-Factor) EW Index. The rebalancing frequency of the strategies is quarterly subject to buffer, turnover control and liquidity capping rules. One-way Turnover is half the sum of absolute deviations of individual weights between the end of a quarter and the beginning of the following quarter. One-way Turnover is annualised for each quarter by multiplying it by 4. The transaction cost estimates use the spread estimates of Closing Quoted spread. The reported transaction cost estimates are the difference between the annualised gross and net returns. Net returns are obtained after accounting for transaction costs at each quarterly rebalancing by multiplying the change in weight of each stock (including stock deletions and additions) between the final weight before rebalancing and the optimal weights after rebalancing by half spread. Days to trade (DTT) refers to the time it takes to trade an initial investment given the stock weights at every quarterly rebalancing date, assuming 10% of the average daily dollar traded volume can be traded. The nominal amount considered for the initial investment equals USD 1bn and is deflated back in time in line with the return of the cap-weighted market-index. The average 95th percentile of DTT is reported which is the time series average of the cross-sectional 95th percentile of DTT at each quarterly rebalancing. For relative calculations we use the Cap-weighted benchmark without accounting for transaction costs (i.e. using its gross returns). Weighted average market capitalisation (in USD millions) is obtained for each quarter by summing the product of the weight of the stock at each quarterly rebalancing by its corresponding total market capitalisation, and then a time-series average across quarters is calculated and reported. The Amihud measure for a given stock is calculated each day by dividing the absolute value of the stock´s return by the traded volume. The Amihud measure is multiplied by 10^9 for ease of presentation. The quarterly average is then obtained by averaging the daily values for which the Amihud measure exists. The weighted average Amihud measure is then obtained for each quarter and a time series average across all quarters are calculated and reported. The Weighted average spread is obtained for each quarter by summing the product of the weight of the stock at each quarterly rebalancing by its corresponding spread estimate – the Closing Quoted spread and then a time-series average across quarters is calculated and reported. Extreme liquidity measures are obtained by sorting the universe by the corresponding liquidity measure and the measure corresponding to the stock that represents the 95% cumulative weight in the corresponding smart beta strategy is reported. Data Source: CRSP, Scientific Beta.

Developed ex USA Short-Term 31-Dec-2004 to 31-Dec-2014	SciBeta Developed ex USA Indices							
	Broad Cap-Weighted		Efficient Minimum Volatility		Maximum Deconcentration		Multi-Beta Multi-Strategy 4-Factor EW	
Panel A – Implementation Costs								
One-Way Turnover	4.36%		32.55%		30.41%		38.98%	
Transaction Costs	0.02%		0.16%		0.16%		0.19%	
Days to Trade (95 %ile)	0.54		2.59		2.20		3.12	
Panel B – Performance Measures								
	Gross	Net	Gross	Net	Gross	Net	Gross	Net
Annualised Returns	5.42%	5.40%	8.00%	7.84%	6.49%	6.33%	7.61%	7.42%
Relative Returns	-	-0.02%	2.58%	2.42%	1.07%	0.91%	2.19%	2.00%
Sharpe Ratio	0.21	0.21	0.43	0.42	0.27	0.26	0.37	0.36
Information Ratio	-	-1.69	0.56	0.52	0.47	0.39	0.63	0.58
Panel C– Liquidity Measures								
	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)
Market Capitalisation	43,901	2,149	14,680	1,074	12,574	985	11,757	980
Spread	0.24%	0.61%	0.39%	0.85%	0.36%	0.89%	0.40%	0.92%
Amihud Measure	6.88	11.40	11.47	18.66	14.05	25.20	13.03	24.35

3. Analysing Smart Beta Strategies

Table 12: Transaction Costs, Performance and Liquidity Measures of Popular Smart Beta Strategies (USA Long Term Track Records (LTTR) – Short Term – 10 Years)
The time period of analysis is 31-Dec-2004 to 31-Dec-2014. All statistics are annualised and daily total returns in USD are used for this analysis. The strategies considered for this analysis are the SciBeta USA LTTR Efficient Minimum Volatility Index, the SciBeta USA LTTR Maximum Deconcentration Index and the SciBeta USA LTTR Multi-Beta Multi-Strategy (4-Factor) EW Index. The rebalancing frequency of the strategies is quarterly subject to buffer, turnover control and liquidity capping rules. One-way Turnover is half the sum of absolute deviations of individual weights between the end of a quarter and the beginning of the following quarter. One-way Turnover is annualised for each quarter by multiplying it by 4. The transaction cost estimates use the spread estimates according to the year of the rebalancing – Range-Based spread until 1993 and Closing Quoted spread from 1993 onwards. The reported transaction cost estimates are the difference between the annualised gross and net returns. Net returns are obtained after accounting for transaction costs at each quarterly rebalancing by multiplying the change in weight of each stock (including stock deletions and additions) between the final weight before rebalancing and the optimal weights after rebalancing by half spread. Days to trade (DTT) refers to the time it takes to trade an initial investment given the stock weights at every quarterly rebalancing date, assuming 10% of the average daily dollar traded volume can be traded. The nominal amount considered for the initial investment equals USD 1bn and is deflated back in time in line with the return of the cap-weighted market-index. The average 95th percentile of DTT is reported which is the time series average of the cross-sectional 95th percentile of DTT at each quarterly rebalancing. For relative calculations, we use the Cap-weighted benchmark for a given stock selection without accounting for transaction costs (i.e. using its gross returns). The Weighted average market capitalisation (in USD millions) is obtained for each quarter by summing the product of the weight of the stock at each quarterly rebalancing by its corresponding total market capitalisation, and then a time-series average across quarters is calculated and reported. The Amihud measure for a given stock is calculated each day by dividing the absolute value of the stock´s return by the traded volume. The Amihud measure is multiplied by 10^9 for ease of presentation. The quarterly average is then obtained by averaging the daily values for which the Amihud measure exists. The weighted average Amihud measure is then obtained for each quarter and a time series average across all quarters are calculated and reported. Weighted average spread is obtained for each quarter by summing the product of the weight of the stock at each quarterly rebalancing by its corresponding spread estimate (Range-Based spread until 1993 and Closing Quoted spread from 1993 onwards) and then a time-series average across quarters is calculated and reported. Extreme liquidity measures are obtained by sorting the universe by the corresponding liquidity measure and the measure corresponding to the stock that represents the 95% cumulative weight in the corresponding smart beta strategy is reported. Data Source: CRSP, Scientific Beta.

USA LTTR Short-Term 31-Dec-2004 to 31-Dec-2014	SciBeta USA LTTR Indices							
	Broad Cap-Weighted		Efficient Minimum Volatility		Maximum Deconcentration		Multi-Beta Multi-Strategy 4-Factor EW	
Panel A – Implementation Costs								
Turnover	2.35%		24.06%		19.77%		28.89%	
Transaction Costs	< 0.01%		0.02%		0.02%		0.02%	
Days to Trade (95 %ile)	0.13		0.55		0.44		0.63	
Panel B – Performance Measures								
	Gross	Net	Gross	Net	Gross	Net	Gross	Net
Annualised Returns	7.71%	7.71%	10.26%	10.24%	9.51%	9.49%	10.26%	10.24%
Relative Returns	-	0.00%	2.55%	2.53%	1.80%	1.78%	2.56%	2.53%
Sharpe Ratio	0.31	0.31	0.49	0.49	0.36	0.36	0.44	0.44
Information Ratio	-	-1.55	0.69	0.69	0.48	0.47	0.85	0.84
Panel C – Liquidity Measures								
	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)	Weighted Average	Extreme Measure (95% Cumulative Weight)
Market Capitalisation	95,513	7,078	31,979	3,822	25,234	3,172	23,322	3,238
Spread	0.06%	0.11%	0.06%	0.13%	0.07%	0.15%	0.07%	0.15%
Amihud Measure	0.07	0.34	0.15	0.59	0.18	0.92	0.18	0.90

Conclusions

Conclusions

A usual warning against smart beta strategies is that they will entail higher replication costs than cap-weighted market indices (Malkiel, 2014). While this is obviously true, the crux of the question is not whether transaction costs are higher but whether, after accounting for such costs, there are any benefits in terms of net returns.

A reasonable expectation from an investor's perspective is that providers should disclose the level of transaction costs generated by their strategies so as to allow for information on net returns. However, providers typically fail to make explicit adjustments for transaction costs and satisfy themselves by reporting gross returns, leaving it to other market participants to figure out what the transaction costs are like.

This paper set out to apply methods for explicit cost measurement, and has drawn conclusions on smart beta strategies. It draws on a methodology that can be easily replicated with low data needs and few computational burdens, thus allowing transaction cost adjustments to be implemented with relative ease.

Our empirical analysis leads to several important conclusions.

In terms of the reliability of low frequency measures that are easy to implement, there is substantial literature suggesting that such measures are highly correlated with high frequency cost measures. Our assessment indeed confirms that low frequency measures reliably capture the level of costs. In particular, we show that our measures capture information content of transaction costs (effective spreads) reported by trading venues in compliance with Rule 605 regulations. They also align well with effective spreads extracted from high frequency trade and quote data (TAQ). Compared to estimates from high frequency data, our cost measures are, however, somewhat conservative in that they tend to slightly overestimate cost levels. This means that any conclusions about the viability of smart beta strategies in the face of transaction costs will also tend to be on the conservative side.

In terms of replication costs estimates that result from applying our methodology, we obtain several interesting findings.

First, we find that conclusions about transaction cost levels and strategy implementation challenges are heavily dependent on the stock universe used. While it is common to see broad brush statements about the investability hurdles of particular smart beta strategies, our results provide clear evidence that conclusions heavily depend on the universe under consideration. Our results on generic strategies show that cost metrics and investability metrics differ tremendously across universes. For example, for portfolios built from the top 250 stocks by market cap, we obtain a Days to Trade measure of 3.56 days compared to 2.06 for the cap-weighted portfolios in the same universe. Moreover, the estimate of average annualised transaction costs is 0.13% for the equal-weighted portfolios compared to 0.04% for the cap-weighted portfolios in the same stock universe. When looking at portfolios formed from the broad universe (the top 3,000 stocks by market cap), we get strikingly different results. The Days to Trade measure reaches more than 100 for equal-weighted portfolios compared to about 10 for cap-weighted portfolios. Estimated transaction costs are 0.38% for equal-weighted portfolios compared to 0.05% for cap-weighted portfolios. Thus an equal weighting strategy indeed

Conclusions

looks extremely challenging to implement for the broad universe, but implementation measures are rather well-behaved for the large-cap universe. Given such differences, it makes little sense to make statements about the investability of any given strategy per se without considering the universe it is implemented for.

Second, our analysis provides evidence of the usefulness of practical implementation rules. Our results suggest that whether or not smart beta strategies face implementation hurdles depends on the set of implementation rules that have been included in the design. We test available index strategies by comparing them to stylised portfolios that omit the implementation rules applied in practice. Our results suggest that smart beta strategies may indeed appear as if they are challenging to implement when abstracting from commonly used implementation rules, but applying these rules leads to different conclusions. For example, we report results for a minimum volatility strategy before applying implementation rules and compare this to the same strategy after such rules have been incorporated. We show that estimated annualised transaction costs change from 0.38% to only 0.18% and investability measures such as Days to Trade go from 3.14 to 2.19 when applying practical investability rules. Perhaps more importantly, amounts traded in any stock relative to its market cap weight decline drastically from a trading multiple of 15 to a multiple of around 1. Applying common sense implementation rules thus reduces transaction costs and limits any stress on available trading volume.

Third, we find that for the set of indices included in our analysis, smart beta performance benefits largely survive transaction costs. When looking at commonly used smart beta indices that are built on liquid universes and integrate implementation rules, the impact of transaction costs on returns is small, far from cancelling out the relative return benefits over cap-weighted indices. Transaction costs are an order of magnitude smaller than relative returns, meaning that net relative returns do not differ materially from gross relative returns. For the three strategies we consider, namely a Minimum Volatility, Maximum Deconcentration and a Multi Factor index, we find that average annualised transaction costs over the 42-year period are between 0.13% to 0.18%, while gross returns relative to the cap-weighted index range from 2.38% to 3.93%.

To summarise the cost conclusions from a practical perspective, here we assess the cost of switching from a cap-weighted portfolio to smart-beta portfolios, and add the resulting rebalancing costs. Table 13 and Table 14 provide further comparison between the stretched transition (where one spreads out the switch to smart beta over various days), and the non-stretched transition from cap-weighted portfolios to smart beta portfolios. It can be seen that the stretching the transition over a period improves the Days to Trade, but the returns remain almost the same. The tracking error between the stretched and non-stretched portfolios also remains quite low although they increase in the stretch period. The cost of transition is very small compared to the cost of rebalancing and the total cost is still low compared to the gross returns even after accounting for the transition costs.

Conclusions

Table 13: Comparison of stretched and non-stretched transition from Cap-Weighted portfolio to Smart Beta Portfolio (Long Term - 42 years)
The time period of analysis is 31-Dec-1972 to 31-Dec-2014. The strategies considered for this analysis are the SciBeta USA LTTR Efficient Minimum Volatility Index, the SciBeta USA LTTR Maximum Deconcentration Index and the SciBeta USA LTTR Multi-Beta Multi-Strategy (4-Factor) EW Index. All statistics reported in Panel A are quarterly estimates and are averaged across all quarters. Results of three types of scenarios are estimated and presented – i) The switch from Cap-Weighted portfolio to Smart Beta portfolio happens completely on the day of rebalancing (1-day Transition); ii) The switch from Cap-Weighted portfolio to Smart Beta portfolio happens equally distributed across 10-days (10-day Transition i.e. assuming only one-tenth of the portfolio switches every day for 10 days); iii) The switch from Cap-Weighted portfolio to Smart Beta portfolio happens equally distributed across 20 days (20-day Transition, i.e. assuming only one-twentieth of the portfolio switches every day for 20 days). Days to Trade (DTT) is reported as a time-series average of the cross-sectional 95th percentile of DTT at each quarterly rebalancing. Tracking Error of stretched transition (both 10-days and 20-days) over non-stretched transition is computed quarterly and average is reported. Difference in Gross Returns is computed quarterly between stretched (both 10-days and 20-days) transition and non-stretched transition. All statistics reported in Panel B are annualised. It compares costs of all three smart beta strategies. Assuming a 10-year investment period, the Annualised Cost of Transition from Cap-Weighted Index is computed as one-tenth of the immediate transition (a semi-absolute difference between weights of smart beta strategies and Cap-Weighted index multiplied by the average weighted spread and averaged across all quarters). Annualised Cost of Rebalancing is the average difference between annualised gross and net returns. Total Annualised Cost is sum of transition and rebalancing costs. Data Source: CRSP and Scientific Beta.

USA LTTR Long-Term 31-Dec-1972 to 31-Dec-2014	Transition	Efficient Minimum Volatility	Maximum Deconcentration	Multi-Beta Multi-Strategy 4-Factor EW
Panel A: Transition from Cap-Weighted Index (statistics for transition quarter)				
Days to Trade (95%ile)	Non-stretched	1.72	1.98	2.64
	Stretched 10-days	0.17	0.20	0.26
	Stretched 20-days	0.09	0.10	0.13
Tracking Error	Non-stretched	-	-	-
	Stretched 10-days	0.08%	0.08%	0.09%
	Stretched 20-days	0.12%	0.11%	0.12%
Difference in Gross Returns by Stretching	Non-stretched	-	-	-
	Stretched 10-days	0.00%	0.01%	0.00%
	Stretched 20-days	-0.01%	0.00%	-0.01%
Panel B: Cost Comparison				
Annualised Cost of Transition from Cap-Weighted (assuming a 10-year investment period)		0.02%	0.02%	0.03%
Annualised Cost of Rebalancing		0.18%	0.13%	0.17%
Total Annualised Cost		0.20%	0.15%	0.20%

The difference in results between Table 13 and Table 14 points to another interesting finding in our study. Indeed, we see a strong time trend in transaction costs. Results discussed so far in the conclusion are based on a long-term history of 42 years. However, transaction costs have declined dramatically as market microstructure changed and competition among trading venues increased. To shed some light on the differences across time, we analysed sub-periods that correspond to different market microstructure regimes. In particular we distinguish between an earlier period when the tick size was 1/8th of a dollar and a subsequent period after decimalisation (a tick size of 1/100th of a dollar) was introduced. Obviously, the reduced tick size allows for lower spreads to be quoted. We find that, in the earlier period, annualised transaction costs ranged from 0.14% to 0.22% for the three index strategies. During the latter period, our cost estimates are only 0.03% to 0.04%. In a nutshell, not only have transaction costs been low enough historically to allow smart beta performance benefits to be extracted in practice, but this hurdle has also shrunk much further over the more recent period.

Conclusions

Table 14: Comparison of stretched and non-stretched transition from Cap-Weighted portfolio to Smart Beta Portfolio (Short Term - 10 years)
The time period of analysis is 31-Dec-2004 to 31-Dec-2014. The strategies considered for this analysis are the SciBeta USA LTTR Efficient Minimum Volatility Index, the SciBeta USA LTTR Maximum Deconcentration Index and the SciBeta USA LTTR Multi-Beta Multi-Strategy (4-Factor) EW Index. All statistics reported in Panel A are quarterly estimates and are averaged across all quarters. Results of three types of scenarios are estimated and presented – i) The switch from Cap-Weighted portfolio to Smart Beta portfolio happens completely on the day of rebalancing (1-day Transition); ii) The switch from Cap-Weighted portfolio to Smart Beta portfolio happens equally distributed across 10-days (10-day Transition i.e. assuming only one-tenth of the portfolio switches every day for 10 days); iii) The switch from Cap-Weighted portfolio to Smart Beta portfolio happens equally distributed across 20 days (20-day Transition, i.e. assuming only one-twentieth of the portfolio switches every day for 20 days). Days to Trade (DTT) is reported as a time-series average of the cross-sectional 95th percentile of DTT at each quarterly rebalancing. Tracking Error of stretched transition (both 10-days and 20-days) over non-stretched transition is computed quarterly and average is reported. Difference in Gross Returns is computed quarterly between stretched (both 10-days and 20-days) transition and non-stretched transition. All statistics reported in Panel B are annualised. It compares costs of all three smart beta strategies. Assuming a 10-year investment period, the Annualised Cost of Transition from Cap-Weighted Index is computed as one-tenth of the immediate transition (a semi-absolute difference between weights of smart beta strategies and Cap-Weighted index multiplied by the average weighted spread and averaged across all quarters). Annualised Cost of Rebalancing is the average difference between annualised gross and net returns. Total Annualised Cost is sum of transition and rebalancing costs. Data Source: CRSP and Scientific Beta.

USA LTTR Short-Term 31-Dec-2004 to 31-Dec-2014	Transition	Efficient Minimum Volatility	Maximum Deconcentration	Multi-Beta Multi-Strategy 4-Factor EW
Panel A: Transition from Cap-Weighted Index (statistics for transition quarter)				
Days to Trade (95%ile)	Non-stretched	0.46	0.40	0.50
	Stretched 10-days	0.05	0.04	0.05
	Stretched 20-days	0.02	0.02	0.03
Tracking Error	Non-stretched	-	-	-
	Stretched 10-days	0.07%	0.06%	0.07%
	Stretched 20-days	0.10%	0.09%	0.11%
Difference in Gross Returns by Stretching	Non-stretched	-	-	-
	Stretched 10-days	-0.01%	0.01%	0.00%
	Stretched 20-days	-0.01%	0.01%	0.00%
Panel B: Cost Comparison				
Annualised Cost of Transition from Cap-Weighted (assuming a 10-year investment period)		<0.01%	<0.01%	<0.01%
Annualised Cost of Rebalancing		0.02%	0.02%	0.02%
Total Annualised Cost		0.02%	0.02%	0.02%

The results in this paper provide an important contribution to the analysis of smart beta strategies from a practical perspective. Indeed, the state of affairs in the evaluation of smart beta strategy performance is far from satisfying. On the one hand, strategy providers do not commonly report the transaction cost estimates of their strategies, and performance evaluation often relies on simulated gross returns. On the other hand, the discussion of cost issues more often than not remains at the level of blanket criticism aimed at certain strategies, without considering the universe or the implementation rules that are used. Our results provide an explicit estimate of costs applied to a range of strategies and show the impact of using different implementation rules or stock universes. Importantly, given the transparent methodology and benign data needs, our replication cost analysis is straightforward and can be easily applied to other strategies.

Conclusions

References

References

- Abdi, F., and A. Ranaldo. 2017. A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices. Working Paper, University of St.Gallen.
- Amenc, N., F. Ducoulombier, F. Goltz and S. Sivasubramanian. 2016. Scientific Beta Multi Smart Factor Indices: A Double Diversification Approach to Factor Investing. ERI Scientific Beta White Paper.
- Amenc, N., F. Goltz, L. Martellini and P. Retkowsky. 2011. Efficient Indexation: An Alternative to Capt-Weighted Indices. *Journal of Investment Management* 9(4): 1-23.
- Amihud, Y. 2002. Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. *Journal of Financial Markets* 5: 31-56.
- Arnott, R. D., J. Hsu and P. Moore. 2005. Fundamental Indexation. *Financial Analysts Journal* 61(2): 83-99.
- Asness, C. S., T. J. Moskowitz and L. H. Pedersen. 2013. Value and Momentum Everywhere. *Journal of Finance* 68(3): 929-985.
- Auer, B. and F. Schuhmacher. 2015. Liquid Betting against Beta in Dow Jones Industrial Average Stocks. *Financial Analysts Journal* 71(6): 30-43.
- Avramov, D., T. Chordia and A. Goyal. 2006. Liquidity and Autocorrelations in Individual Stock Returns. *Journal of Finance* 61(5): 2365-2394.
- Bennett, P. and L. Wei. 2006. Market Structure, Fragmentation, and Market Quality. *Journal of Financial Markets* 9(1): 49-78.
- Blitz, D. C. and P. van Vliet. 2007. The Volatility Effect. *Journal of Portfolio Management* 4(1): 102-113.
- Boehmer, E. 2005. Dimensions of Execution Quality: Recent Evidence for US Equity Markets. *Journal of Financial Economics* 78(3): 553-582.
- Brunnermeier, M. 2009. Deciphering the Liquidity and Credit Crunch 2007-2008. *Journal of Economic Perspectives* 23(1): 77-100.
- Chan, L. K. C., J. Karceski and J. Lakonishok. 1999. On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model. *Review of Financial Studies* 12(5): 937-974.
- Choueifaty, Y. and Y. Coignard. 2008. Toward Maximum Diversification. *Journal of Portfolio Management* 35(1): 40-51.
- Chow, T., J. Hsu, V. Kalesnik and B. Little. 2011. A Survey of Alternative Equity Index Strategies. *Financial Analysts Journal* 67(5): 37-57.
- Christensen, B. J. and N. R. Phabhala. 1998. The Relation Between Implied and Realized Volatility. *Journal of Financial Economics* 50: 125-150.
- Chung, K. H. and H. Zhang. 2014. A Simple Approximation of Intraday Spreads with Daily Data. *Journal of Financial Markets* 17: 94-120.
- Clarke, R., H. de Silva and S. Thorley. 2006. Minimum-Variance Portfolios in the U.S. Equity Market. *Journal of Portfolio Management* 33(1): 10-24.

References

- Connor, G., L. R. Goldberg and R. A. Korajczyk. 2010. Transaction Costs and Liquidity Risk. In Portfolio Risk Analysis. Princeton, NJ: Princeton University Press.
- Corwin, S. A. and P. Schultz. 2012. A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices. *Journal of Finance* 67(2): 719-760.
- Dash, S. and K. Loggie. 2008. Equal Weighting – Five Years Later. Standard & Poor's Publication.
- De Groot, W., J. Huij and W. Zhou. 2011. Another Look at Trading Costs and Short-Term Reversal Profits. *Journal of Banking & Finance* 36(2): 371-382.
- DeMiguel, V., L. Garlappi, J. Nogales and R. Uppal. 2009a. A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms. *Management Science* 55(5): 798-812.
- DeMiguel, V., L. Garlappi and R. Uppal. 2009b. Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies* 22(5): 1915-1953.
- Fama, E. F. and K. R. French. 2004. The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives* 18(3): 25-46.
- Fong, K., C. Holden, and C. Trzcinka. 2014. What Are The Best Liquidity Proxies For Global Research? Working paper.
- Frazzini, A., R. Israel and T. J. Moskowitz. 2015. Trading Costs of Asset Pricing Anomalies. Working paper.
- Frazzini, A. and L. H. Pedersen. 2014. Betting Against Beta. *Journal of Financial Economics* 111: 1-25.
- Garleanu, N. and L. H. Pedersen. 2011. Margin-based Asset Pricing and Deviations from the Law of One Price. *Review of Financial Studies* 24(6): 1980-2022.
- Garman, M. B. and M. J. Klass. 1980. On the Estimation of Security Price Volatilities from Historical Data. *The Journal of Business* 53(1): 67-78.
- Geiger, H. and J. Plagge. 2007. Minimum Variance Indices. Frankfurt: Deutsche Börse AG.
- Giglio, S. and K. Shue. 2014. No News is News: Do Markets Underreact to Nothing? *Review of Financial Studies* 27(12): 3389-3440.
- Goltz, F. and A. Lodh. 2015. Scientific Beta Efficient Minimum Volatility Indices. ERI Scientific Beta White Paper.
- Gonzalez, N., S. Sivasubramanian and S. Ye. 2015. Scientific Beta Maximum Deconcentration Indices. ERI Scientific Beta White Paper.
- Goyenko, R. Y., C. W. Holden and C. A. Trzcinka. 2009. Do Liquidity Measures Measure Liquidity? *Journal of Financial Economics* 92: 153-181.
- Hasbrouck, J. 2007. Empirical Market Microstructure. New York, NY: Oxford University Press.

References

- Hasbrouck, J. 2009. Trading Costs and Returns for U.S. Equities: Estimating Effective Costs from Daily Data. *Journal of Finance* 64(3): 1445-1477.
- Holden, C. W. 2009. New Low-frequency Spread Measures. *Journal of Financial Markets* 12(4): 778-813.
- Huang, R. D. and H. R. Stoll. 1996. Dealer Versus Auction Markets: A Paired Comparison of Execution Costs on NASDAQ and the NYSE. *Journal of Financial Economics* 41: 313-357.
- ITG Peer Analysis. 2016. Global Cost Review. Accessed April 20. Available at: http://analyticsincubator.itginc.com/data/itg-peer-analysis-global-cost-review/.
- Jagannathan, R. and T. Ma. 2003. Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *Journal of Finance* 58(4): 1651-1684.
- Karnaukh, N., A. Ranaldo and P. Söderlind. 2015. Understanding FX Liquidity. *Review of Financial Studies* 28(11): 3073-3108.
- Keim, D.B. and A. Madhavan. 1997. Transactions Costs and Investment Style: An Inter-Exchange Analysis of Institutional Equity Trades. *The Journal of Financial Economics* 46: 265-292.
- Lee, C. M. C. and M. J. Ready. 1991. Inferring Trade Directions from Intraday Data. *Journal of Finance* 46(2): 733-746.
- Leland, H. E. 1999. Optimal Portfolio Management with Transaction Costs and Capital Gains Taxes. Working paper, UC Berkeley.
- Lesmond, D., J. Ogden and C. Trzcinka. 1999. A New Estimate of Transaction Costs. *Review of Financial Studies* 12(5): 1113-1141.
- Maillard, S., T. Roncalli and J. Teïletche. 2010. The Properties of Equally Weighted Risk Contribution Portfolios. *Journal of Portfolio Management* 36(4): 60-70.
- Malkiel, B. G. 2014. Is Smart Beta Really Smart? *The Journal of Portfolio Management* 40(5): 127-134.
- Marshall, B. R., N.H. Nguyen and N. Visaltanachoti. 2012. Commodity Liquidity Measurement and Transaction Costs. *Review of Financial Studies* 25(2): 599-638.
- McLean, D. and J. Pontiff. 2016. Does Academic Publication Destroy Stock Return Predictability? *Journal of Finance*, Forthcoming. Available at: http://ssrn.com/abstract=2156623.
- Nagel, S. 2012. Evaporating Liquidity. *Review of Financial Studies* 25(7): 2005-2039.
- Nielsen, F. and R. Aylursubramanian. 2008. Far from the Madding Crowd – Volatility Efficient Indices. MSCI Barra Research Insights.
- Novy-Marx, R. and M. Velikov. 2016. A Taxonomy of Anomalies and their Trading Costs. *Review of Financial Studies* 29(1): 104-147.
- O'Hara, M. and M. Ye. 2011. Is Market Fragmentation Harming Market Quality? *Journal of Financial Economics* 100(3): 459-474.

References

- Parkinson, M. 1980. The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business* 53(1): 61-65.
- Platen, E. and R. Rendek. 2010. Approximating the Numeraire Portfolio by Naive Diversification. *Journal of Asset Management* 13(1): 34-50.
- Poullaouec, T. 2008. Things to Consider when Investing in Minimum Variance Strategies. State Street Global Advisors.
- Roll, R. 1977. A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory. *Journal of Financial Economics* 4(2): 129-176.
- Roll, R. 1984. A Simple Implicit Measure of the Effective Bid-ask Spread in an Efficient Market. *Journal of Finance* 39(4): 1127-1139.
- Schestag, R., P. Schuster and M. Uhrig-Homburg. 2016. Measuring Liquidity in Bond Markets. *Review of Financial Studies* 29(5): 1170-1219.
- Schwartz, T. 2000. How to Beat the S&P 500 with Portfolio Optimization. Working paper, DePaul University.
- Sharpe, W. F. 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance* 19(3): 425-442.
- West, J. 2007. Price-Indifferent Indexes: The Equal-Weighted Alternative. Research Affiliates. Available at: https://www.researchaffiliates.com/en_us/publications/articles/412_Price-Indifferent_Indexes_The_Equal-Weighted_Alternative.html.
- Zhao, X. and K. H. Chung. 2007. Information Disclosure and Market Quality: The Effect of SEC Rule 605 on Trading Costs. *Journal of Financial and Quantitative Analysis* 42(03): 657-682.

References

About Amundi ETF, Indexing & Smart Beta

About Amundi ETF, Indexing & Smart Beta

The Amundi ETF, Indexing and Smart Beta business line is one of Amundi group's strategic business areas and totalizes 64bn € AuM[1].

Built on strong commitments on cost efficiency, innovation and transparency, the Amundi ETF platform is the 5th largest ETF provider in Europe[2] with 100 ETFs and more than 450 listings across Europe.

On Indexing and Smart Beta, innovation and customization are at the core of the client-oriented approach. The objective is to provide investors with robust, flexible and highly cost efficient solutions, leveraging on Amundi pricing power and extensive resources, including first class research capabilities in SRI and Factor investing.

1 - Source: Amundi ETF, Indexing & Smart Beta as of 31/12/2016
2 - Source: *Deutsche Bank European Monthly ETF market review*, December 2016

About EDHEC-Risk Institute

About EDHEC-Risk Institute

Founded in 1906, EDHEC is one of the foremost international business schools. Accredited by the three main international academic organisations, EQUIS, AACSB, and Association of MBAs, EDHEC has for a number of years been pursuing a strategy of international excellence that led it to set up EDHEC-Risk Institute in 2001. This institute now boasts a team of close to 50 permanent professors, engineers and support staff, as well as 38 research associates from the financial industry and affiliate professors.

The Choice of Asset Allocation and Risk Management and the Need for Investment Solutions

EDHEC-Risk has structured all of its research work around asset allocation and risk management. This strategic choice is applied to all of the Institute's research programmes, whether they involve proposing new methods of strategic allocation, which integrate the alternative class; taking extreme risks into account in portfolio construction; studying the usefulness of derivatives in implementing asset-liability management approaches; or orienting the concept of dynamic "core-satellite" investment management in the framework of absolute return or target-date funds. EDHEC-Risk Institute has also developed an ambitious portfolio of research and educational initiatives in the domain of investment solutions for institutional and individual investors.

Academic Excellence and Industry Relevance

In an attempt to ensure that the research it carries out is truly applicable, EDHEC has implemented a dual validation system for the work of EDHEC-Risk. All research work must be part of a research programme, the relevance and goals of which have been validated from both an academic and a business viewpoint by the Institute's advisory board. This board is made up of internationally recognised researchers, the Institute's business partners, and representatives of major international institutional investors. Management of the research programmes respects a rigorous validation process, which guarantees the scientific quality and the operational usefulness of the programmes.

Six research programmes have been conducted by the centre to date:

- Asset allocation and alternative diversification
- Performance and risk reporting
- Indices and benchmarking
- Non-financial risks, regulation and innovations
- Asset allocation and derivative instruments
- ALM and asset allocation solutions

These programmes receive the support of a large number of financial companies. The results of the research programmes are disseminated through the EDHEC-Risk locations in Singapore, which was established at the invitation of the Monetary Authority of Singapore (MAS); the City of London in the United Kingdom; Nice and Paris in France.

EDHEC-Risk has developed a close partnership with a small number of sponsors within the framework of research chairs or major research projects:

- ETF and Passive Investment Strategies, *in partnership with Amundi ETF*
- Regulation and Institutional Investment, *in partnership with AXA Investment Managers*
- Asset-Liability Management and Institutional Investment Management, *in partnership with BNP Paribas Investment Partners*
- New Frontiers in Risk Assessment and Performance Reporting, *in partnership with CACEIS*
- Exploring the Commodity Futures Risk Premium: Implications for Asset Allocation and Regulation, *in partnership with CME Group*

About EDHEC-Risk Institute

- Asset-Liability Management Techniques for Sovereign Wealth Fund Management, *in partnership with Deutsche Bank*
- The Benefits of Volatility Derivatives in Equity Portfolio Management, *in partnership with Eurex*
- Structured Products and Derivative Instruments, *sponsored by the French Banking Federation (FBF)*
- Optimising Bond Portfolios, *in partnership with the French Central Bank (BDF Gestion)*
- Risk Allocation Solutions, *in partnership with Lyxor Asset Management*
- Infrastructure Equity Investment Management and Benchmarking, *in partnership with Meridiam and Campbell Lutyens*
- Risk Allocation Framework for Goal-Driven Investing Strategies, *in partnership with Merrill Lynch Wealth Management*
- Investment and Governance Characteristics of Infrastructure Debt Investments, *in partnership with Natixis*
- Advanced Modelling for Alternative Investments, *in partnership with Société Générale Prime Services (Newedge)*
- Advanced Investment Solutions for Liability Hedging for Inflation Risk, *in partnership with Ontario Teachers' Pension Plan*
- Active Allocation to Smart Factor Indices, *in partnership with Rothschild & Cie*
- Solvency II, *in partnership with Russell Investments*
- Structured Equity Investment Strategies for Long-Term Asian Investors, *in partnership with Société Générale Corporate & Investment Banking*

The philosophy of the Institute is to validate its work by publication in international academic journals, as well as to make it available to the sector through its position papers, published studies, and global conferences.

To ensure the distribution of its research to the industry, EDHEC-Risk also provides professionals with access to its website, www.edhec-risk.com, which is entirely devoted to international risk and asset management research. The website, which has more than 70,000 regular visitors, is aimed at professionals who wish to benefit from EDHEC-Risk's analysis and expertise in the area of applied portfolio management research. Its quarterly newsletter is distributed to more than 200,000 readers.

EDHEC-Risk Institute: Key Figures, 2014-2015

Number of permanent staff	48
Number of research associates & affiliate professors	36
Overall budget	€6,500,000
External financing	€7,025,695
Nbr of conference delegates	1,087
Nbr of participants at research seminars and executive education seminars	1,465

About EDHEC-Risk Institute

Research for Business

The Institute's activities have also given rise to executive education and research service offshoots. EDHEC-Risk's executive education programmes help investment professionals to upgrade their skills with advanced risk and asset management training across traditional and alternative classes. In partnership with CFA Institute, it has developed advanced seminars based on its research which are available to CFA charterholders and have been taking place since 2008 in New York, Singapore and London.

In 2012, EDHEC-Risk Institute signed two strategic partnership agreements with the Operations Research and Financial Engineering department of Princeton University to set up a joint research programme in the area of asset-liability management for institutions and individuals, and with Yale School of Management to set up joint certified executive training courses in North America and Europe in the area of risk and investment management.

As part of its policy of transferring know-how to the industry, in 2013 EDHEC-Risk Institute also set up ERI Scientific Beta. ERI Scientific Beta is an original initiative which aims to favour the adoption of the latest advances in smart beta design and implementation by the whole investment industry. Its academic origin provides the foundation for its strategy: offer, in the best economic conditions possible, the smart beta solutions that are most proven scientifically with full transparency in both the methods and the associated risks.

EDHEC-Risk Institute Publications and Position Papers (2014-2017)

EDHEC-Risk Institute Publications (2014-2017)

2016

- Amenc, N., F. Goltz, V. Le Sourd. Investor Perceptions about Smart Beta ETFs (August).
- Giron, K., L. Martellini and V. Milhau Multi-Dimensional Risk and Performance Analysis for Equity Portfolios (July).
- Maeso, J.M., L. Martellini. Factor Investing and Risk Allocation: From Traditional to Alternative Risk Premia Harvesting (June).
- Amenc, N., F. Goltz, V. Le Sourd, A. Lodh and S. Sivasubramanian. The EDHEC European ETF Survey 2015 (February).
- Martellini, L. Mass Customisation versus Mass Production in Investment Management (January).

2015

- Blanc-Brude, F., M. Hasan and T. Whittaker. Cash Flow Dynamics of Private Infrastructure Project Debt (November).
- Amenc, N., G. Coqueret, and L. Martellini. Active Allocation to Smart Factor Indices (July).
- Martellini, L., and V. Milhau. Factor Investing: A Welfare Improving New Investment Paradigm or Yet Another Marketing Fad? (July).
- Goltz, F., and V. Le Sourd. Investor Interest in and Requirements for Smart Beta ETFs (April).
- Amenc, N., F. Goltz, V. Le Sourd and A. Lodh. Alternative Equity Beta Investing: A Survey (March).
- Amenc, N., K. Gautam, F. Goltz, N. Gonzalez, and J.P Schade. Accounting for Geographic Exposure in Performance and Risk Reporting for Equity Portfolios (March).
- Amenc, N., F. Ducoulombier, F. Goltz, V. Le Sourd, A. Lodh and E. Shirbini. The EDHEC European Survey 2014 (March).
- Deguest, R., L. Martellini, V. Milhau, A. Suri and H. Wang. Introducing a Comprehensive Risk Allocation Framework for Goals-Based Wealth Management (March).
- Blanc-Brude, F., and M. Hasan. The Valuation of Privately-Held Infrastructure Equity Investments (January).

2014

- Coqueret, G., R. Deguest, L. Martellini, and V. Milhau. Equity Portfolios with Improved Liability-Hedging Benefits (December).
- Blanc-Brude, F., and D. Makovsek. How Much Construction Risk do Sponsors take in Project Finance. (August).
- Loh, L., and S. Stoyanov. The Impact of Risk Controls and Strategy-Specific Risk Diversification on Extreme Risk (August).

EDHEC-Risk Institute Publications (2014-2017)

- Blanc-Brude, F., and F. Ducoulombier. Superannuation v2.0 (July).
- Loh, L., and S. Stoyanov. Tail Risk of Smart Beta Portfolios: An Extreme Value Theory Approach (July).
- Foulquier, P. M. Arouri and A. Le Maistre. P. A Proposal for an Interest Rate Dampener for Solvency II to Manage Pro-Cyclical Effects and Improve Asset-Liability Management (June).
- Amenc, N., R. Deguest, F. Goltz, A. Lodh, L. Martellini and E.Schirbini. Risk Allocation, Factor Investing and Smart Beta: Reconciling Innovations in Equity Portfolio Construction (June).
- Martellini, L., V. Milhau and A. Tarelli. Towards Conditional Risk Parity – Improving Risk Budgeting Techniques in Changing Economic Environments (April).
- Amenc, N., and F. Ducoulombier. Index Transparency – A Survey of European Investors Perceptions, Needs and Expectations (March).
- Ducoulombier, F., F. Goltz, V. Le Sourd, and A. Lodh. The EDHEC European ETF Survey 2013 (March).
- Badaoui, S., Deguest, R., L. Martellini and V. Milhau. Dynamic Liability-Driven Investing Strategies: The Emergence of a New Investment Paradigm for Pension Funds? (February).
- Deguest, R., and L. Martellini. Improved Risk Reporting with Factor-Based Diversification Measures (February).
- Loh, L., and S. Stoyanov. Tail Risk of Equity Market Indices: An Extreme Value Theory Approach (February).

EDHEC-Risk Institute Publications (2014-2017)

EDHEC-Risk Institute Position Papers (2014-2017)

2016

- O'Kane, D. Initial Margin for Non-Centrally Cleared OTC Derivatives (June).

2014

- Blanc-Brude, F. Benchmarking Long-Term Investment in Infrastructure: Objectives, Roadmap and Recent Progress (June).

For more information, please contact:
Carolyn Essid on +33 493 187 824
or by e-mail to: carolyn.essid@edhec-risk.com

EDHEC-Risk Institute
393 promenade des Anglais
BP 3116 - 06202 Nice Cedex 3
France
Tel: +33 (0)4 93 18 78 24

EDHEC Risk Institute—Europe
10 Fleet Place, Ludgate
London EC4M 7RB
United Kingdom
Tel: +44 207 871 6740

EDHEC Risk Institute—Asia
1 George Street
#07-02
Singapore 049145
Tel: +65 6438 0030

EDHEC Risk Institute—France
16-18 rue du 4 septembre
75002 Paris
France
Tel: +33 (0)1 53 32 76 30

www.edhec-risk.com