Deutsche Bank
Markets Research

Europe

Derivatives Strategy
Derivatives Spotlight

Date
20 April 2017

Caio Natividade
caio.natividade@db.com

Silvia Stanescu
silvia.stanescu@db.com

Vivek Anand
vivek-v.anand@db.com

Paul Ward, Ph.D
paul.ward@db.com

Simon Carter
simon-d.carter@db.com

Pam Finelli
pam.finelli@db.com

Spyros Mesomeris, Ph.D
spyros.mesomeris@db.com

+44 20 754 52198

Volatility Risk Premium: New Dimensions

Today's *Derivatives Spotlight* delves into systematic options research. It is the first in a series of collaborative reports between our derivatives and quantitative research teams that aim to systematically identify and capture value across global volatility markets.

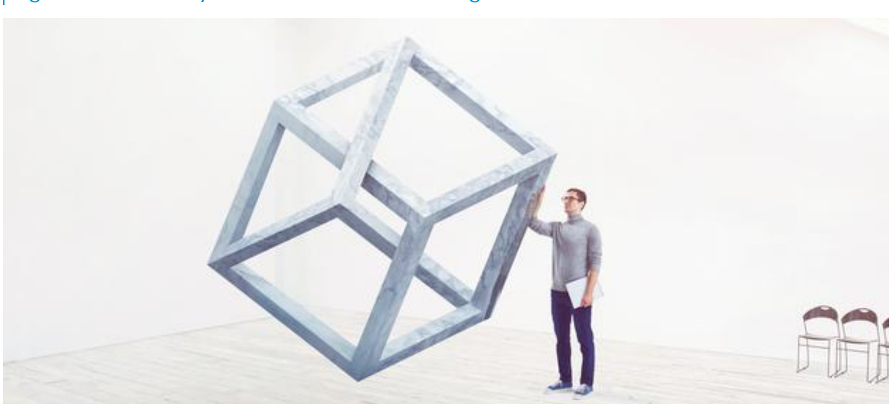
This edition zooms into the volatility risk premia (VRP), one of the key sources of return in options markets. VRP strategies are popular across the investor community, but suffer from structural shortcomings. This report looks to improve on those.

Going beyond traditional methods, we introduce a P-distribution that best represents our projected future returns and associated probabilities, based on their drivers. Other topics are also highlighted as we construct our P-distribution, namely a new multivariate volatility risk factor model, our Global Sentiment Indicator, and the treatment of event-based versus non-event based returns.

We formulate a strategy which should improve the way in which the VRP is harnessed. It utilizes alternative delta hedging methods and timing.

Risk Statement: while this report does not explicitly recommend specific options, we note that there are risks to trading derivatives. The loss from long options positions is limited to the net premium paid, but the loss from short option positions can be unlimited.

Figure 1: Volatility Risk Premium: Thinking Outside The Box



Source: Getty Images.

1. Introduction

Options trading is popular among discretionary and systematic investors, as it serves multiple purposes in institutional investment portfolios. Today's *Derivatives Spotlight* is the first in a series of collaborative reports between our derivatives and quantitative research teams that aim to systematically identify and capture value across global volatility markets.

While we assume some technical knowledge, the report is aimed at both novice and experienced investors alike. For ease of navigation we outline each of the sections below:

- Section 2 looks at the primary drivers of options returns – our starting point. We look at both delta hedged and "naked" option return streams, and find that the former is largely driven by volatility while the latter is associated with underlying asset returns. The patterns exhibited in the first principal components point to the existence of a volatility risk premium (VRP) - the focus of this report.
- Section 3 expands on the VRP and highlights standard measures to tame strategy drawdowns. We highlight the benefits and shortcomings of cross-market replication (3.1), market timing (3.2) and buying tail options (3.3), and allude to new approaches - namely the P-distribution and alternative delta hedging.
- Section 4 introduces our approach to modeling the P-distribution, the first new route to improve how we capture value from options markets. It represents our best estimate of future returns and return probabilities based on their drivers. Section 4.2 describes how we estimate the mean, and Section 4.3 provides in-depth coverage of our variance estimates. These are based on a multivariate risk factor model (4.3.1 and Appendix I), a regime-dependent multivariate GARCH model (4.3.2 and Appendix II) and event-specific volatility (4.3.3). The higher moments of our P-distribution are described in Section 4.4, while Section 4.5 provides a thorough assessment of how our forecasts compare to a number of benchmarks.
- Section 5 describes alternative delta hedging – the second avenue we take as we seek to improve VRP strategy returns. We introduce 5 methods from industry and academia: grid-search (5.1), moving averages (5.2), break outs (5.3), expected returns (5.4) and a modified version of Whalley & Wilmott (5.5). All are tested under both simulated data (5.6) and real market data (5.7).
- Section 6 focuses on building our new VRP strategy. We first fine tune the signal (6.1), then add alternative delta hedging to the mix (6.2). Section 6.3 deals with timing via our Global Sentiment Indicator, long-term volatility reversals, and near-term volatility changes. Section 6.4 compares methods, while Section 6.5 concludes with an application to the S&P 500, Eurostoxx 50 and Nikkei 225.
- Section 7 concludes the report, while Appendix I covers our multi-variate risk factor model and Appendix II provides a review of our Global Sentiment Indicator. Our Bibliography is presented at the end.

We believe the VRP is a generic premia, whose dynamics and investment implications extend beyond multiple markets and asset classes. As such, our analysis covers all 4 asset classes: equity indices, commodities, currencies and global Treasuries. We look at options in the S&P 500, Eurostoxx 50, Nikkei 225, Bovespa, EUR/USD, USD/JPY AUD/USD, USD/BRL, Gold, WTI, Corn, Copper, 10Y US Treasuries, 10Y Bunds and 10Y JGBs. While their liquidity characteristics are distinctly different, each should represent a regional hub inside the asset class and allows us to understand the volatility premia more holistically.

Most of our market data is sourced internally. We proxy variance swaps using strike-weighted baskets of European vanilla options. In some sections we use a purer basket comprised of puts and calls whose strikes expand up to 2.5 standard deviations away from spot. In other sections we use a more realistic basket based on delta strikes (10-delta, 25-delta and ATM). All options are rebalanced at expiry. All results prior to Section 6 omit costs; we only apply them at the end as the VRP strategy gets finalized.

As is often the case in quantitative research, some results are not as encouraging as we initially hoped. Other results turned out to be better than expected. In all, we believe that the final recommendations - based on alternative delta hedging and timing - should improve the way investors harness this rich source of risk-adjusted returns.

2. The Primary Factors

Understanding what drives returns is key to investing successfully in any market. To the quant investor, the learning process often involves running principal component analysis (PCA) on the returns of all available market instruments and interpreting what the output – the core factors – means.

As we delved into the world of options data, our approach was no different: collect instrument returns and run the PCA. For any given market, we took a set of relevant calls and puts – 10- and 25-delta, and ATMF – across relevant maturities – 1, 3, 6 and 12 months – and calculated 2 historical long-only return streams[1]: delta hedged and un-hedged. Therefore, each of the 15 options markets outlined earlier had 48 return series: 24 for delta-hedged options, and 24 for "naked" option returns. We used up to 12 years of returns data, where available.

The PCA output[2] set the base for our approach to factor investing in options markets. The key findings from the aggregate of all PCAs were that:

- Spot moves are the chief return driver of "naked" option contracts, explaining circa 63% of total variations in each market, on average. This is no surprise; the delta exposure, and hence the delta return, dwarfs all other "Greeks" in the tenors evaluated.
- Changes in volatility were the main source of delta-hedged option returns, causing 54% of the variations on average. This is the main source of risk for the volatility trader.
- Changes in volatility – implied (delta) vols, in particular – are also the second driver of naked option returns, accounting for 17% of the variations in the sample data. Lower delta options are more sensitive to this factor.
- The shape of the volatility smile and the steepness of the term structure explain an extra 20% of the variations in delta-hedged option returns. In other words, the volatility trader cannot ignore skew and slope dynamics either.
- The returns from delta-hedged options are more diverse – the first 2 principal components explain 65% of the variations on average, versus 80% in the case of naked option returns.

Figure 2: Correlations between PCs and asset returns, implied and realized volatility

	PC1 "naked" returns			PC1 delta hedged returns			PC2 "naked" returns		
	A. Ret.	RV	IV	A. Ret.	RV	IV	A. Ret.	RV	IV
E/$	0.9	-0.3	-0.3	-0.1	0.7	0.8	-0.2	0.4	0.7
$/Y	0.9	-0.1	-0.2	-0.1	0.8	0.8	0.2	0.4	0.6
A/$	0.9	-0.5	-0.5	-0.4	0.8	0.8	-0.2	0.6	0.6
$/BR	0.9	0.6	0.6	0.4	0.9	0.8	0.1	0.5	0.6
S&P 500	0.9	-0.5	-0.7	-0.6	0.8	0.8	0.0	0.3	0.2
SXE	0.8	-0.4	-0.6	-0.4	0.7	0.6	-0.2	0.3	0.4
Nik	0.9	-0.4	-0.5	-0.5	0.8	0.8	-0.1	0.2	0.4
Bov	0.9	-0.4	-0.6	-0.5	0.7	0.7	-0.3	0.3	0.6
Gold	0.9	-0.2	0.4	0.0	0.8	0.8	0.0	0.4	0.6
WTI	0.9	-0.3	-0.4	-0.4	0.8	0.8	-0.3	0.4	0.6
Corn	0.9	-0.1	0.3	-0.2	0.6	0.6	0.0	0.2	0.3
Cop	1.0	-0.6	-0.6	-0.6	0.8	0.8	-0.6	0.6	0.6
UST	0.9	0.0	0.2	0.1	0.4	0.7	-0.3	-0.1	0.3
Bun	0.8	0.1	0.2	-0.1	0.6	0.6	-0.2	0.5	0.4
JGB	0.7	0.0	-0.1	-0.2	0.6	0.4	-0.6	0.2	0.6

E/$: EUR/USD, $/Y: USD/JPY, A/$: AUD/USD, $/BR: USD/BRL, SXE: Eurostoxx 50, Nik: Nikkei 225, Bov: Bovespa, Cop: Copper, Bun: 10Y Bunds. A.Ret.: annualized returns, RV: realized volatility (1Y lookback, 3-month half-life), IV: implied volatility. Source: Deutsche Bank

[1] We used Black & Scholes pricing as these are vanilla European options. We built each surface using SSVI interpolation and extrapolation. SSVI stands for Surface Stochastic Volatility Inspired. We used the approach introduced in Gatheral and Jacquier [2013].

[2] We used up to 11 years of daily returns, fixed investment notionals in each option (so that return volatilities were more similar) and estimated the principal components using a correlation matrix instead of the covariance matrix.

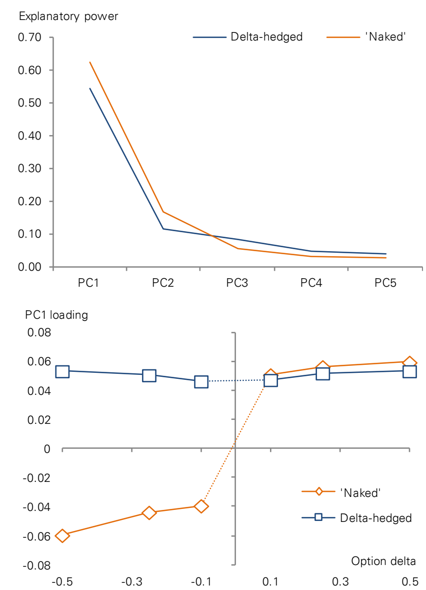
Figure 3 – an aggregate of the results for all 15 markets – display our conclusions visually. The chart at the top shows how dominant the first principal components are in each type of return streams, while the bottom chart shows how they load at each point of the volatility smile, after aggregating by expiries. The quasi-homogenous loadings per strike in the delta-hedged streams, and the strong but inverse looking loadings in the "naked" streams support the argument that the primary factor of returns heavily affects all options in the surface. As we correlated the historical PC1 values with more *tangible* variables, we found an 88% average correlation[3] with underlying asset returns (PC1 of "naked" options) and a 72% average correlation with realized volatility (PC1 of delta-hedged option returns). Confirming the points above, spot and volatility are the main drivers of "naked" and delta-hedged option returns.[4]

Next, we analyse the historical developments of these first principal components over time in search for potential biases. *Bias* is what we ultimately look for when scrolling through markets with our "quant lenses", as it lies at the core of most systematic strategies. The stronger the pattern, and the more that it can be validated, the more likely that we should succeed in capturing it.[5]

Take the PC1 of "naked" option returns, for instance. The high correlation to underlying spot markets suggests that the *alpha capture* will come from forecasting future spot values, just as we do in our delta-one portfolios. The profit of a "naked" option trade depends on where the asset will be relative to its break-even at expiry, and the latter reflects a market-implied probability that the option will expire in-the-money. As such, we should try to estimate *our own* probabilities that the option will expire ITM, and emit signals based on where the probability spread is most significant. We call this the *probability risk premium* (PRP), and it will be the subject of future research.

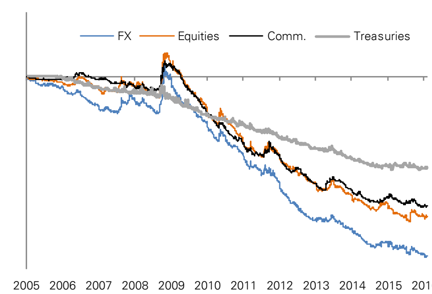
Today's report focuses on the PC1 of delta-hedged returns, whose historical values are aggregated by asset class in Figure 4. It has a distinct characteristic: *regardless of the underlying asset, or asset class, this principal component trends down over time*. This obvious bias suggests that buying delta-hedged options is a losing proposition over time. And since theory argues that excess returns should be zero if the "true" volatility is embedded into the option price[6], we conclude that the bias originates from implied volatility being generally above what is realized. This bias suggests that selling delta-hedged options is a profitable trade over time, and we call it the *volatility risk premium* (VRP). That said, the "hick-ups" in the series also show that buying options can be attractive once in a while, which means our capturing process should be dynamic; we should not just blindly sell options.

Figure 3: Explanatory power and factor loading by principal component of option returns



Up to 12 years of data on option returns. See section above for which assets and strikes are included. Source: Deutsche Bank

Figure 4: Cumulative standardised PC1 values per asset class



The standardization went as follows: we took daily returns of each option for a given market and divided by its 1-year volatility. We then estimated the PC1 loadings using a correlation matrix with 5 years of daily data, and calculated the most recent PC1 value from the loadings. We repeated the exercise daily. We then averaged the PC1 values for each asset class and plotted the cumulative values. Source: Deutsche Bank

3 Monthly non-overlapping returns, 11 years of data.

4 We did not extend the study to cover, for instance, the drivers of spot returns and of spot volatility. We recommend Natividade et al [2013] (Section 7) and Natividade et al [2014] (Section 2) for the former; for the latter, we recommend Mccormick and Natividade [2008], Natividade [2008] and Saravelos and Grover [2017] in FX and Corradi et al [2006] and Engle and Rangel [2005] in equities.

5 Two examples from delta-one markets illustrate our argument: the PC1 of commodity futures returns, and the PC3 of a given yield curve. The former exhibits positive and stable autocorrelation over time, which validates trend following strategies in commodity markets. The latter represents *curvature* and often exhibits strong mean-reverting properties, which validates systematic butterfly strategies in liquid IRS markets.

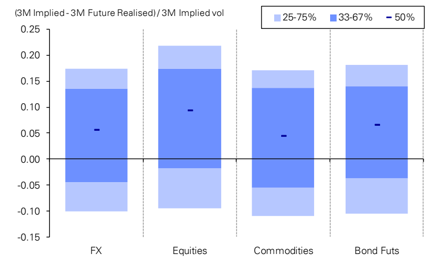
6 Assuming other theoretical conditions hold.

3. Exploring the Volatility Risk Premium (VRP)

The VRP is normally defined as compensation for systematic risk, price pressures from investor supply and demand, and biased forecasts of future empirical volatility.[7] We measure it as *the distance between current implied and future empirical volatilities*. Therefore, it is no surprise that this metric is not only positive[8] in every asset class, but more expressive in markets with heavier natural demand for crash insurance – such as equity indices.

Figure 5 shows the historical VRP distribution for the aforementioned pool of 15 assets, bucketed by asset class, using the 3-month horizon as reference and expressing the VRP as a ratio to implied volatility for better comparison across markets. While built differently, it correlates strongly to the inverse of the PC1 of delta-hedged returns as described in Section 2. And as we group all asset classes together, we find that the VRP has been stable and positive in 61% of the instances. It has also been more strongly pronounced in equities, which reinforces the demand argument outlined above.

Figure 5: Historical VRP distributions by asset class



Up to 12 years of data. We use 15 markets (see Section 1).
Source: Deutsche Bank

VRP strategies primarily aim to capture that difference between implied and realized volatilities. Volatility and variance swaps are the standard instrument of choice for the VRP harvester, though a similar P&L profile can also be achieved through selling baskets of vanilla options with frequent delta hedging and capital allocation that is inversely related to the strike level.[9] We opt for such baskets as they give us more delta hedging freedom, which will come handy later. In this section, we use baskets comprised of 1-month options with 12 strikes that are equally distanced and range from -2.5 to +2.5 standard deviations away from spot.[10] All options were delta hedged daily and rolled at maturity, and we omitted costs prior to the final section of this report.

Figure 6 shows the cumulative returns of our baskets above, standardized such that they are equally volatile and therefore easier to visualize. As expected, they are also similar to the inverse of the PC1s from Figure 4, even though we used a different weighting scheme to group the individual options.

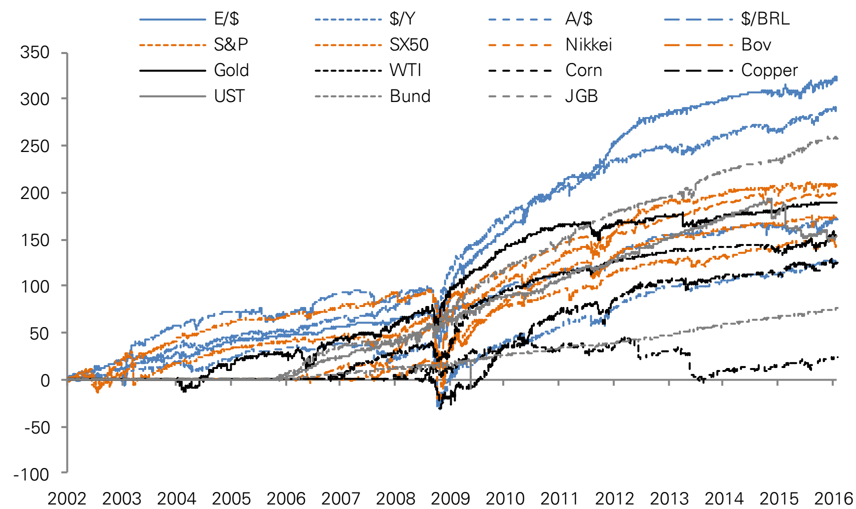
[7] See Ilmanen [2011] for a general review. Other authors who defended similar arguments include Carr and Wu [2009], Benzoni et al [2010] and Broadie and Johannes [2009].

[8] That said, the VRP is far less pronounced in single stock equities than in equity indices and other asset classes. As argued by Cosemans [2011] and Valenzuela [2014], such difference is attributed to positive correlation risk premium - how index implied correlations over-estimate future realised correlations. Our research also validates their argument – see Prasad et al. [2016].

[9] This is, in fact, the way a market maker replicates a variance swap, as it generates a stable gamma profile across different spot levels. For a recent reference, see Derman and Miller [2016].

[10] We use the standard deviation of monthly returns as a measure. Some of these strikes may not be tradable, but suit this exercise as we are looking for a good variance swap replication that is also computationally efficient.

Figure 6: Cumulative, volatility-adjusted pre-cost 1-month VRP basket returns per asset



Labels: E/$: EUR/USD, $/Y: USD/JPY, A/$: AUD/USD, $/BRL: USD/BRL, S&P: S&P 500, SX50: Eurostoxx 50, Nikkei: Nikkei 225, Bov: IBOVESPA, UST: 10Y US Treasuries, Bund: 10Y Bunds, JGB: 10Y JGBs. Source: Deutsche Bank

In general terms, the investor who sells an option is selling some form of insurance. By selling the whole smile, we are selling market insurance much more broadly; significant moves in any direction are likely to hurt us. Most assets involved fall faster than they rise, and therefore "significant moves" typically coincide with *bearish markets*. In 2012[11] we launched the DB Global Sentiment Indicator (GSI), a variable that captures market risk appetite and whose details are described in Appendix II. Figure 7 shows how it relates to our VRP baskets: irrespective of the asset class, falling VRP returns are associated with rising GSI levels, indicating *higher risk aversion*.

[11] See Natividade and Chen [2012b].

In essence, the VRP strategy is *pro-cyclical*; it underperforms when markets are risk averse – most commonly when macroeconomic volatility is high and output is contracting. Needless to say, this is a shortcoming. It means that VRP strategies correlate positively with the static – equity, credit – premia that dominate institutional portfolios and, as such, that they also underperform when core portfolios are suffering. Adding insult to injury, VRP drawdowns are notable for their magnitude; the 4 months following the Lehman Brothers announcement in Sep-08 wiped out the previous 6 years of returns in a typical tradable S&P 500 VRP strategy[12], and 8 years of returns in short USD and EUR tradable swaption strategies.[13]

There are 3 popular methods that seek to improve the profile of VRP returns: cross-market replication, factor timing and buying tail options.

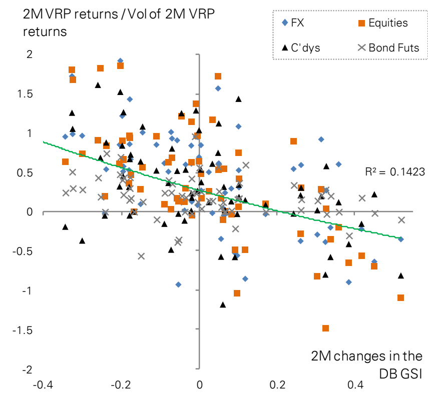
3.1 Cross-market Replication

Cross-market replication is popular in the quantitative investment community, especially when dealing with generic delta-one factors such as Momentum and Carry. Since the number of independent trades influences the Sharpe ratio of any strategy, the quant investor will seek to maximize breadth by replicating the same signal across a large number of markets.[14]

While attractive, the idea has a shortcoming when applied on the VRP. As is often the case with other *base* factor strategies applied to a limited asset pool, the VRP signal has limited breadth because it captures a "global" driver that explains the majority of moves in all instruments available.[15] Investor aversion – the abstract risk *whose* premium is captured by short volatility strategies – is common across asset classes, leading to a strong link between option returns even in markets where spot returns are not highly correlated. Figure 8 shows exactly that: the VRP return correlations in the bottom triangle are, on average, stronger than the spot return correlations in the upper triangle.

Cross-market replication improves the return profile – the expected shortfall shrinks by almost one third in our backtest[16] – but it is not enough on its own, as we seek a better overall improvement.

Figure 7: 2M cross-asset VRP returns versus DB's Global Sentiment Indicator



The 2-month VRP returns comprise all 15 markets, aggregated by asset class. The reader may be intrigued as to why we chose the 2-month time window. Shorter windows produced a sparser relationship (partly related to Footnote 15), while longer windows had too few datapoints. Source: Deutsche Bank

[12] We use a standard, tradable index for this comparison: the DB Equity US Volatility Carry Index (DBGLSVEU Index on Bloomberg). This strategy sells 2M and 3M straddles on the S&P 500, delta-hedged daily. Costs are included.

[13] We use a standard, tradable index for this comparison: DB ImpAct (DBIP3BE Index and DBIP3BU Index on Bloomberg). This strategy sells 3M10Y USD and EUR swaption straddles every week, delta-hedged at a given frequency. Costs are included.

[14] This follows from the Law of Fundamental Active Management, where $Sharpe \approx IC \times \sqrt{breadth}$. IC, the information coefficient, represents signal forecast accuracy and breadth stands for the number of independent investment decisions. See Grinold & Kahn [1999].

[15] See Baz et al [2015], and Natividade et al [2016a]. These references also show that in order to preserve signal entropy, base factor strategies are best implemented in time series form - in other words, without committing to an equal number of long and short trades.

[16] We use the same VRP replicating basket described earlier in the section and data since 2002. Shortfalls are defined as the average of monthly returns equal to or lower than the 5th percentile of monthly returns in the sample. The average shortfall is estimated by averaging the shortfalls of all 15 VRP strategies, weighted by volatility, whereas the shortfall of the average is estimated by grouping all VRP strategies (weighted by volatility) and then calculating the monthly shortfall.

Figure 8: Correlations between underlying markets – VRP returns (bottom triangle) and spot returns (top triangle)

	E/$	Y/$	A/$	BRL/$	S&P	SX50	Nikkei	Bov	Gold	WTI	Corn	Copper	UST	Bund	JGB
E/$		0.23	0.72	0.50	0.39	0.29	0.24	0.48	0.43	0.51	0.37	0.51	-0.14	-0.32	0.03
Y/$	0.55		0.17	-0.09	-0.24	-0.19	-0.51	-0.10	0.37	-0.04	-0.02	-0.20	0.47	0.32	0.41
A/$	0.73	0.63		0.76	0.60	0.49	0.39	0.72	0.46	0.60	0.55	0.66	-0.02	-0.31	-0.12
BRL/$	0.65	0.55	0.85		0.47	0.42	0.35	0.62	0.41	0.60	0.46	0.64	-0.20	-0.37	-0.16
S&P	0.65	0.52	0.82	0.79		0.91	0.78	0.71	-0.08	0.36	0.33	0.58	-0.20	-0.40	-0.27
SX50	0.53	0.46	0.67	0.71	0.71		0.74	0.70	-0.04	0.21	0.25	0.48	-0.09	-0.35	-0.24
Nikkei	0.61	0.67	0.68	0.66	0.72	0.52		0.51	-0.24	0.22	0.12	0.47	-0.25	-0.30	-0.40
Bov	0.52	0.60	0.66	0.65	0.67	0.51	0.52		0.26	0.51	0.34	0.55	-0.13	-0.42	-0.29
Gold	0.45	0.45	0.45	0.46	0.53	0.37	0.40	0.62		0.30	0.29	0.26	0.09	-0.09	0.09
WTI	0.58	0.51	0.52	0.54	0.62	0.39	0.61	0.57	0.42		0.24	0.65	-0.28	-0.50	-0.21
Corn	0.41	0.32	0.41	0.34	0.40	0.30	0.50	0.39	0.31	0.41		0.42	0.03	-0.18	-0.03
Copper	0.42	0.31	0.61	0.60	0.71	0.51	0.49	0.46	0.48	0.47	0.26		-0.23	-0.43	-0.12
UST	0.31	0.18	0.17	0.18	0.17	0.36	0.21	0.12	0.08	0.18	0.12	0.04		0.64	0.41
Bund	0.37	0.23	0.17	0.15	0.17	0.33	0.13	0.08	0.20	0.16	0.16	0.02	0.60		0.56
JGB	-0.03	0.04	-0.20	-0.25	-0.24	-0.06	-0.04	-0.28	0.10	-0.14	-0.06	-0.13	0.10	0.25	

Correlation of 2-month non-overlapping returns, Note that we use FX/USD and not USD/FX to make VRP and spot market correlations more comparable. 11-year history. Source: Deutsche Bank

3.2 Market Timing

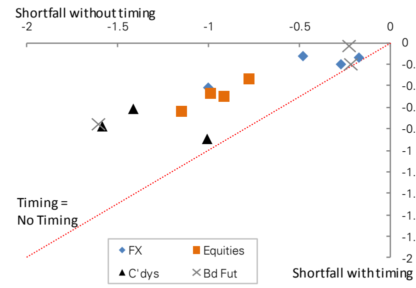
Timing the VRP strategy is also popular and, if the backtests serve as an indication, it can be successful as well. The premise is that if there is a strong relationship between strategy returns and global risk appetite, and if we can *time* the latter, we should be able to reduce the drawdowns in the former.

To test it, we created a linearly continuous leverage mechanism to the VRP strategies above using the same GSI as before. The higher the GSI, the less we sold of our options baskets and vice-versa; in other words, we de-levered during periods of risk aversion and levered up when markets were calm. The leverage ratio is defined as $L_{t,t+h} = 2 - GSI_t \times 2$, and is therefore constrained such that $L \in [0,2]$. The leverage changes every 2 weeks, and it has an average of 1.[17]

[17] Unlike in prior applications, we apply the GSI leverage on a bi-weekly basis instead of daily. We evaluated the sensitivity to changes in leveraging frequency and found that "high" frequencies (daily) and very low frequencies (bi-monthly) yield worse results. Daily rebalancing incurs more signal noise and factor reversal risk: *when the GSI is at extreme levels, next-day strategy returns are at greater risk of going against the timing indicator than next-week or next-month returns.* We also knew that once we attach transaction costs to this exercise, daily rebalancing becomes unfeasible. On the other hand, very low frequencies, such as bi-monthly, are not adaptive enough. Of the frequencies tested, the highest risk-adjusted returns came when using monthly rebalancing. We chose bi-weekly for illustration purposes alone. These results do not include costs - we leave that to later in this report.

History suggests this is an attractive idea. Figure 9 plots the effect from leveraging our strategy according to the GSI. The X-axis represents the original backtested monthly shortfall[18] in the VRP strategy for each of our 15 markets, labeled by asset class, while the Y-axis shows the new shortfall after applying the technique above. The sample covers 11 years of data. While we only capture a handful of crisis periods, the results are encouraging in that most of the effect comes from taming our drawdowns. They also concur with our positive experience timing risk-sensitive, pro-cyclical signals in the past – notably the FX Carry trade, as shown in Appendix II.

Figure 9: Backtested monthly shortfalls – cross-market VRP strategy



We define shortfall as the average monthly return using months that qualified as 5th percentile of worst months in the sample for each strategy. Source: Deutsche Bank

3.3 Buying Tail Options

Practitioners have also been fond of buying "tail options" to improve VRP strategy drawdowns. The method is straight forward: sell volatility at the nearby strikes as per original strategy but buy volatility at distant strikes so as to flip the gamma exposure in the event that spot moves significantly.[19] Ideally, if the spot market has fallen aggressively, and volatility has therefore risen, our net gamma exposure should have gone from short to long. In principle, therefore, this should improve VRP return drawdowns.

In order to test the idea, we modified our option baskets slightly. Instead of only using 1-month options as earlier in this section, we used 4 expiries: 1, 3, 6 and 12 months. We also changed our strikes; instead of 12 *fixed-percent-distance* strikes equally spaced in steps of ½ of a standard deviation, we used 6 *fixed-delta strikes* with distance defined in delta steps: 10-delta, 25-delta and ATMF puts and calls. These revised maturities allow us to capture different theta decay profiles, while using delta strikes allows us to better compare cross-maturity strikes.

The goal of this backtest was to evaluate the sensitivity of VRP strategy returns to using one set of specific strikes as hedge. In other words, we went short one unit of every option described above except for those with a pre-specified delta. In the latter case, we would go long one unit (of each) instead. The benchmark in this case is to go short all options, with no protection. As before, we delta hedged daily and did not include costs.

[18] We define *shortfall* as the average monthly return using months that qualified as 5th percentile of worst months in the sample for each strategy. In other words, the average return in the 7 worst months in the sample.

[19] A variation of this idea is to only sell delta-hedged straddles. In this case, aggressive spot moves will also have a smaller adverse impact on the VRP strategy because much of the straddle gamma will have been eroded.

The results here are also encouraging. Our Sharpe ratios rose and, more importantly, our drawdowns improved once we balanced our short exposures with a few long positions.

But the improvement also depended on which strikes we chose to go long; in most instances, it is only the low delta puts – the *tail* – that helped. Figure 10 shows the ratio of average monthly shortfall to monthly returns in each of our 15 markets, expressed as a difference to the ratio achieved using the benchmark strategy. The more negative the number, the better – it means that a typical drawdown erased less return months.

Take the Nikkei as an example. In the benchmark strategy, the average monthly shortfall erased 25 months' worth of average monthly returns. If we modified the strategy so as to be long 10-delta puts in the 4 maturities while still short all the other options in the surface[20], the ratio would have fallen from 25 to 18, an improvement of 7 months (hence -7 in the table). Parallel examples can also be made with low delta puts in most other markets, particularly where asset returns exhibit strong negative skew.[21] While one may argue that the return profile should improve further once we optimize our choices for theta decay efficiency[22], these preliminary results are already encouraging.

Figure 10: Shortfall / average returns (VRP with tail options) minus shortfall / average returns (VRP benchmark)

	10p	25p	St'dle	25c	10c
E/$	0.44	0.0	0.3	0.0	0.0
Y/$*	0.7	1.3	-1.5	-0.4	0.5
A/$	-4.9	-6.8	1.1	53.6	5.9
BRL/$*	-2.7	-2.7	3.5	18.7	3.0
S&P	-2.4	-0.1	0.1	3.1	1.2
SX50	-2.4	-3.2	-0.7	9.6	3.2
Nikkei	-7.0	-6.2	-4.0	34.1	8.8
Bov	0.1	9.6	1.3	-2.0	-0.9
Gold	-1.7	0.2	2.8	1.2	-0.3
WTI	-0.4	2.3	0.0	0.1	-0.3
Corn	-9.0	-0.3	-4.9	7.1	23.1
Copper	-2.5	-1.2	-0.6	3.9	1.7
UST	1.7	3.8	3.6	8.1	-4.5
Bund	-0.8	-0.6	3.4	2.1	-0.1
JGB	-0.7	-0.9	8.9	2.0	-0.6

** Flipped USD per unit of currency for better comparison with other markets. Source: Deutsche Bank*

3.4 A New Way of Thinking

These methods are encouraging, but narrow in scope. They focus mostly on controlling strategy drawdowns, and do so by either *timing* risk aversion, or by *hedging* or *diversifying* against it. In all 3 cases, we use prior knowledge and experience to define what risk aversion means and how it manifests itself across markets. We assume, for instance, that risk barometer (the GSI) is accurate, and that only the left tail of asset returns should be of concern to the VRP investor. Further, true instances of risk aversion are scarce - only a handful in our backtest window - which makes our results even more dependent on these priors.

In the remaining sections, we seek to introduce methods that are distinctly different from the ones described above. First, we will not assume that volatility risk premia will always be positive; we will compare the forecasts from our real world distribution – introduced in the next section – with market-implied volatility so as to define the sign and quantity of our VRP positions. Second, we will use our knowledge of the underlying asset returns – direction and trendiness – to delta hedge better. We start with the former in Section 4, as we look at the whole P-distribution instead of asset variance alone. We then move onto delta hedging in Section 5.

[20] For clarity: we went long 10-delta 1M, 3M, 6M and 1Y put options, allocating 1 JPY unit to each. At the same time, we went short 10-delta calls, and 25-delta and ATMF calls and puts in the same maturities, allocating 1 JPY unit to each as well. The capital allocation is therefore 4/24 into long positions, and 20/24 into short positions.

[21] The versions that are long 10-delta puts outperformed except in a few instances - most notably in Treasury futures, a typical safe haven asset, where the version long 10-delta calls outperformed. This reinforces the idea of being long the area of the smile most likely to outperform in the case of risk aversion. That said, a few cases are less intuitive – notably with straddles in JPY/USD and Bovespa calls.

[22] We did not test theta decay optimisation thoroughly, but we tried a rough version that bought only 1Y 10-delta puts (4/24 units) and sold all other options (20/24 units). This version also outperformed the unconstrained VRP strategy, but underperformed the version that also sold 10-delta puts of other expiries.

4. The P-Distribution

As is the case in all derivatives markets, profiting from options trading requires taking a view (of the future) that differs from what is implied by the market price. But options markets carry a unique advantage – their prices contain a far more granular view of the future than futures and forward markets do. By mapping the volatility smile into a risk neutral distribution, one can estimate the market-implied probability that spot will be at any level by horizon date, in addition to estimating how volatile or leptokurtic the market "expects" the returns to be. [23]

This abundance of market-implied views is valuable to the quantitative investor because it increases the number of uncorrelated trades available using the same type of instrument. The investor is no longer constrained to a bullish or bearish position versus the futures price; she can instead trade various options strikes to express higher order views. She can even express a view that is – in principle – indifferent to spot levels, as does the VRP harvester.

The most holistic way to identify option trading opportunities is, in our view, by estimating a distribution of future asset returns and comparing it with the market-implied distribution. In other words, comparing our *P-distribution* with the market's *Q-distribution*. This sets the base for how we will assess value across strikes, expiries and markets. While today's report will only focus on the VRP – differences in the second moment of the two distributions – we will introduce the whole approach as it serves as base for future reports. This section, therefore, introduces our P-distribution.

4.1 P-distribution Basics

The *P-distribution* is our subjective assessment of what future asset returns will look like. Estimating it is a challenge. We have to account for realistic financial market assumptions while also staying computationally efficient. With that in mind, there are 4 aspects to our approach:

- We use our delta-one portfolios to estimate expected returns - the first moment of our P-distribution. Each portfolio contains some additional predictive power about future returns across horizons. The portfolios were built as part of our *Quantcraft* series: Trend, Carry, Value, Sentiment and Macro Factor investing.
- We combine a multivariate *risk factor* model, in addition to a univariate model, to estimate asset variance. This approach utilises both cross-market and idiosyncratic information to forecast the second moment of our P-distribution, and is in line with how quant investment portfolios are built.

[23] As with other derivatives, market-implied views are derived through arbitrage-free relationships that most often reflect the cost of hedging the position and supply-demand dynamics, as opposed to traders' subjective views of the future. This difference is often the source of opportunity.

Specifically to options markets, Carr and Madan [2001] argue that the optimal trade for an investor requires 3 ingredients: (1) her beliefs regarding future outcomes (the P-distribution), (2) her risk preferences (level of risk aversion), and (3) market prices.

- We account for volatility jumps by segregating between event and non-event volatility. The former relates to known economic releases and policy decisions in each country.
- We also account for the influence of global sentiment regimes when calibrating our parameters.
- Finally, we apply an analytical, moment-based approach to estimate the higher moments and ultimately complete our P-distribution. This approach has the advantage of being computationally efficient.

Our returns process is defined as follows:

$$r_{t,i} = \mu_{t,i} + B_i F_t + \varepsilon_{t,i}$$

$$E_{t-1}\left[r_{t,i}\right] = \mu_{t,i}$$

$$Var_{t-1}\left(r_{t,i}\right) = B_i \Omega_{t\|t-1} B_i^T + \Psi_{t|t-1}$$

where:

- $\mu_{t,i}$ is determined exogenously by our delta-one signals. It is non-stochastic but varies over time. It is described in more detail in Section 4.2.
- B_i is a vector of factor loadings: $B_i = (\beta_{1,i}, \beta_{2,i}, \ldots \beta_{K,i})^T$, for i = 1:N assets and k = 1:K factors. F_t is a vector of de-trended[24] observed *risk factor* returns. Both are described in more detail in Section 4.3 and Appendix I.
- $\varepsilon_{t,i} = \sqrt{h_{t,\varepsilon_i}}\, z_{t,i}$ is a disturbance process such that $\varepsilon_{t,i} \sim \mathrm{N}\left(0, h_{t,\varepsilon_i}\right)$.
- Ω_t is a time-varying factor covariance matrix, our *systematic risk*. It is described in more detail in Section 4.3.
- Ψ_t is a time-varying diagonal[25] covariance matrix of error terms. It represents *asset-specific* risk.

4.2 P-Distribution: the Mean

As the P-distribution should represent our "best guess" of future values of spot, it must incorporate all our knowledge about what drives spot returns. So far, our work indicates five categories of drivers – trendiness, valuations, carry, market sentiment and macroeconomic developments. The respective signals are grouped through the portfolios highlighted in Section 4.1.

For any horizon, our forecasted expected return, μ , is computed as follows:

1. The sign of our forecast reflects the average of the standardised weights[26] from each portfolio.
2. The magnitude of our forecast reflects both the magnitude of recent asset returns and a propagation function. We use a variance ratio as the latter.[27]

Mathematically,

[24] A necessary requirement for $E_{t-1}\left[r_{t,i}\right] = \mu_{t,i}$.

[25] We assume the disturbance terms are uncorrelated.

[26] We divide the current weight by its recently volatility. The absolute weights are bounded at 2.

[27] We use an anchored long-term window to estimate the variance ratio term structure.

$$\mu_{t,i} = f\left(\left|r_{t,i}\right|, \left|r_{t-1,i}\right|, \ldots, \left|r_{0,i}\right|\right) \times \sum_{m=1}^{M} \tilde{w}_{t,m,i}$$

where $f(\bullet)$ is an exponential moving average of *absolute* asset returns. $\tilde{w}_{t,m,i}$ is the standardized time-t weight of asset i associated with portfolio Y_m. In other words, our forecast reflects how bullish or bearish our delta-one portfolios are now versus the past, and whether spot is likely to trend or mean-revert over a given horizon.

4.3 P-Distribution: the Variance

The variance is the second step in estimating our P-distribution. It is the main quantity of interest for the VRP harvester and hence a key focus of this paper. Our approach must reflect real-world characteristics, while remaining computationally efficient.

To model Ω_t we apply a multivariate factor constant conditional correlation GARCH model – CCC MGARCH.[28] The building blocks are as follows:

- *Common factor vs idiosyncratic risk*: we discriminate between systematic and asset-specific risk.
- *A multi-regime approach:* we use the aforementioned DB Global Sentiment Indicator - GSI[29] to exogenously incorporate different risk regimes into our volatility model.
- *Event-based jump risk*: we also incorporate the impact of scheduled macro-economic events into our variance forecasts.

4.3.1 Factor model

Factor models are commonplace in the quantitative investment community. They carry two[30] main advantages:

- *Computational efficiency*. In essence, factor models reduce dimensionality. Instead of modeling the variations and co-variations of N assets individually, we model those of K factors instead. By default, $K < N$. Factor co-variances eventually map back into asset co-variances through the latter's sensitivity to each factor.
- *Estimation robustness*. In a world with increasingly integrated markets, a factor model captures the common drivers – the factors – of cross asset

[28] The fact that financial returns are not well described by *iid* normal distributions has been long documented in the literature, starting as early as 1960s – see, for example Mandelbrot [1963]. Volatility clustering – i.e. positively autocorrelated variances – is well established. The seminal papers of Engle [1982] and Bollerslev [1986] enable the modeling of such volatility clustering. Jumps in the price and or/ the volatility of assets (around events) is another well documented phenomenon. Furthermore, multi-regime GARCH models have been advocated in the relatively recent academic literature – see, for example, Haas et al.[2004a, 2004b], Marcucci [2005], Alexander and Lazar [2006, 2009]. All these features have been incorporated in the volatility model we propose and describe below. CCC-MGARCH stands for Constant Conditional Correlation Multivariate Generalised Autoregressive Conditional Heteroskedasticity.

Factor models for modeling the volatility of assets, particularly of large portfolios, have been long advocated in the academic and practitioner finance literature. Here we are combining the approaches proposed in the seminal papers of Engle et al. [1990] on factor- ARCH, and Bollerslev [1990] on the constant conditional correlation GARCH model, later developed to allow for dynamic correlation by a series of authors, including, for example Engle [2002] or Tse and Tsui [2002].

[29] See Natividade and Chen [2012b] for the methodology behind the construction of our Global Sentiment Indicator (GSI). According to this regime indicator, we distinguish between three risk/market sentiment states, namely: risk-seeking, intermediate and risk aversion.

[30] Another, albeit secondary advantage is that a factor model can also incorporate spill-overs between volatilities of different assets.

variations better. As has been argued in the literature[31], the inter-dependence between the various volatilities may be due to an underlying, or a set of underlying common factors. Sections 2 and 3.1 also illustrate this argument; not only the common variations in delta-hedged returns were highly linked to volatility, but those volatilities were highly correlated across different markets as well.

We model systematic risk through K = 11 global *risk factors*, summarized in Figure 11. We divide them into 2 categories: *macro* and *investable*. The former covers asset sensitivity to broad macroeconomic dynamics – namely growth and inflation. The latter represents directly investable portfolios, either *static* (long-only) or *dynamic* (long-short). Our model is based on the Arbitrage Pricing Theory (APT), an extension of the CAPM where multiple factors are used. Asset exposure to each factor is estimated using a stepwise robust regression, a method that tackles collinearity and outlier effects. Our model covers 80 assets across multiple asset classes and is updated daily. It is covered in detail in Appendix I.

Figure 11: Factors used for Multivariate Variance forecasting

Factor type	Summary
Macro	US inflation (proxied by TIPS vs USTs) Global growth (proxied by equities)
Investable: static	Equities: EM FX: USD/G10, USD/EM Commodities: energy, metals, agriculture Rates: global 10Y Treasuries
Investable: dynamic	Cross-asset Momentum Cross-asset Carry

We use the terms "Macro" and "Investable" as per literature.
Source: Deutsche Bank.

Having outlined the factors, we now relate them to our final asset variance forecast. Recall from Section 4.1 above that the time t variance of asset i is given by:

$$Var_{t-1}(r_{t,i}) = \underbrace{B_i \Omega_{t\|t-1} B_i^T}_{systematic\ risk} + \underbrace{\Psi_{t|t-1}}_{idiosyncratic\ risk}$$

In the context of our factor CCC-MGARCH, we can write Ω_t as the following matrix product:

$$\Omega_t = D_t \Pi_t D_t$$

where $D_t = \underset{1\le k\le K}{diag}\left(\sqrt{h_{t,k}}\right)$is a diagonal matrix of time-varying factor standard deviations and Π_t is a KxK factor correlation matrix. $\Psi_t = \underset{1\le i\le N}{diag}\left(\sqrt{h_{t,\varepsilon_i}}\right)$is a diagonal matrix encapsulating asset specific risk.

: The estimation of our CCC-MGARCH includes the following steps:

- Step1: estimate the factor variance terms (i.e. the elements of D_t) via a GARCH-type model of choice – in our case, a multi-state model where the states are exogenously given by our GSI. This is described in Section 4.3.2.
- Step 2: Repeat Step 1 for the residual variances (i.e. the elements of Ψ_t).
- Step 3: Estimate the correlation matrix Π_t[32]; in the context of a CCC model, $\Pi_t = \Pi$, i.e. correlations are assumed constant through time.

4.3.2 The exogenously-determined multi-state GARCH model

As the seasoned investor is aware, asset volatility follows multiple regimes over time. Capturing regime-shifts quickly enough is key. In our view, the GSI suits that task better than other standard choices such as historical returns alone.

[31] See for example Bauwens et al. [2005] – Chapter 5 in particular and references therein – , Corradi et al. [2006] or Engle and Rangel [2005].

[32] With the sole condition that this is a positive definite matrix, such that the positive definiteness of Ω_t is ensured (i.e. variances are guaranteed to be positive such that volatilities are defined)

The variance of the k-th de-meaned factor $F_{k,}$ conditional on information available at t-1, is given by:

$$\mathrm{var}_{t-1}\left(F_{t,k}\right)=h_{t,k}=\omega_k+\alpha_k F_{t-1,k}^2+\lambda_{k,s_{t-1}}F_{t-1,k}^2+\beta_k h_{t-1,k}$$

The variance terms $h_{t,k}$ (k = 1,2…., 12) will represent the diagonal elements of Ω_t . We also use the same specification for the idiosyncratic variances, namely:

$$\mathrm{var}_{t-1}\left(\varepsilon_{t,i}\right)=h_{t,\varepsilon_i}=\omega_i+\alpha_i\varepsilon_{t-1,i}^2+\lambda_{i,s_{t-1}}\varepsilon_{t-1,i}^2+\beta_i h_{t-1,\varepsilon_i}$$

where s_t is an observable random variable. At any time t, we assume that s_t is equal to either *1*, *2*, or *3*, denoting low, medium or high risk regimes,[33] respectively, depending on the GSI value at t.[34]

For parameter estimation via maximum likelihood, as well as forecasting higher moments and approximate distributions (see Section 4.4), the following, equivalent expression for $h_{t,k}$ will be used:[35]

$$h_{t,k}=\omega_k+\alpha_k F_{t-1,k}^2+\lambda_{1,k}F_{t-1,k}^2 I_{t-1,k}^{(1)}+\lambda_2 F_{t-1,k}^2 I_{t-1,k}^{(2)}+\beta_k h_{t-1,k}$$

where $I_{t,k}^{(i)}$ is an indicator function, taking the value of *1* if s_t = i and zero otherwise, with $i\in\{1,2\}$.

4.3.3 Event versus non-event volatility

It is well established that asset returns can be "jumpy", and we must account for this feature if we want to make our variance forecasts more realistic. While the standard approach involves a compound Poisson process, we opted for an alternative method that targets macroeconomic events – the periods when those jumps are more likely to occur. This bespoke approach is more straightforward and computationally efficient in helping us estimate the P-distribution as a whole.[36] Our approach follows Natividade et al. [2011], which is in turn inspired by Bauwens et al. [2005]. We apply it to both factor and idiosyncratic variances.

---

[33] In Natividade and Chen [2012b], a Gaussian mixture of historical GSI values pointed to 3 as the optimal number of regimes to describe the series.

[34] Identification requires one of the $\lambda_{k,s_{t-1}}$ parameters, with s_{t-1}= {1, 2, 3}, to be equal to zero (or equivalently that α is equal to zero). In the equivalent specification used below, we implicitly assume $\lambda_{k,3}=0$. The same treatment applies to $\lambda_{i,s_{t-1}}$.

[35] A similar expression will be used for the estimation of idiosyncratic risk.

[36] See Cont and Tankov [2004] for a comprehensive review of the parametric treatment of jumps in financial returns in continuous time. Using discrete time modeling, Guegan et al. [2013] propose a GARCH-type model augmented with compound Poisson jumps in the returns. Their proposed approach is thus related to the (continuos-time) jump diffusion model introduced in the seminal work of Merton [1976], and also to the Bernoulli diffusion model (BDM) – see Honore [1998] and Ball and Torous [1983]. The jump-diffusion modeling approach would have the advantage of modeling both (scheduled) events as well as non-(scheduled) event-driven jumps. It would however have the disadvantage is that the estimation of such models is not particularly straightforward, especially in our multivariate setting and given our aim of forecasting cumulative variances and higher moments over various horizons.

Figure 12: Macroeconomic events covered

US	Eurozone	Japan	Australia
FOMC	GE IFO	Balance of Payments: Current Account	Employment Change
Fed Minutes	PMI Composite	Japan Buying Foreign Bonds	GDP
NFP	HICP	Japan Buying Foregin Stocks	CPI
GDP	Unemployment	Foreign Buying Japan Bonds	Retail Sales
ISM Manufacturing	PMI Manufacturing	Foreign Buying Japan Bonds	Labour Price
ISM Services	ECB Policy rate decision	CPI	Policy Rate Decision
CPI	France PMI Manufacturing	GDP	Brazil
Ret. Sales	France PMI Services	Tankan Survey	Policy Rate Decision
Cons. Conf.	Germany PMI Manufacturing	BOJ Policy rate decision	CPI
Home Sales	Germany PMI Services		
House Starts	Italy PMI Manufacturing		
Jobless Claims	Italy PMI Services		

Source: Deutsche Bank; Certain holidays (e.g. London. US) were also treated as events where applicable (i.e. market not closed). Country-specific holidays were also excluded for each asset.

Our approach to forecasting the total variance of asset returns combines the modeling of time-varying volatility from Section 4.3.2 with non-parametric, event-driven jumps. The process involves the following steps:

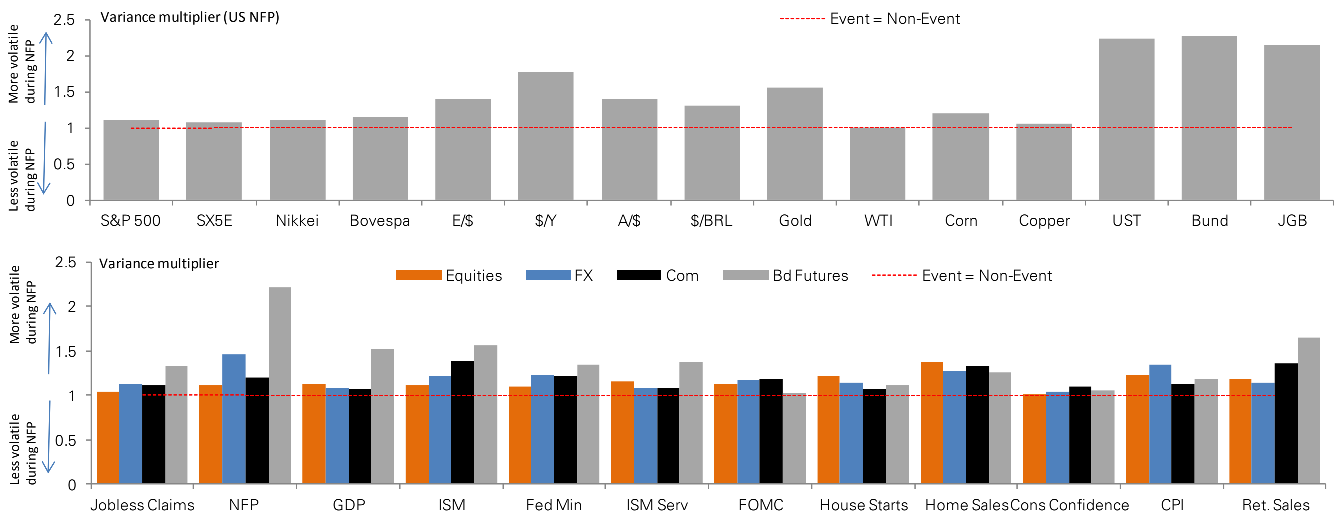
- Step 1: we "clean" the returns data by removing the impact of event-driven jumps. We estimate the impact of a scheduled event by dividing asset variance on the day of the event by asset variance during non-event days. Figure 12 summarizes the events we cover. These ratios will turn into multipliers in the future once we have forecasted "clean" variance.[37] Figure 13 illustrates how significant our calendar events are for a selection of assets; *the fact that the non-holiday events generally have ratios in excess of 1 validates our approach*.[38]
- Step 2: we estimate the parameters of the GARCH process for (clean) variance using the event-adjusted returns series obtained in Step 1.[39]
- Step 3: we compute forward *h*-step ahead "clean" daily variance forecasts based on the parameters estimated in Step 2.
- Step 4: we re-insert the impact of scheduled events by using the ratios from Step 1 as multipliers to the (forward-starting) daily variances from Step 3. Finally, we integrate the forecasts so as to arrive at the final forecast of cumulative, annualized volatility for a given horizon.

[37] Factor models use US events in the modeling of systematic (factor) variance, whereas events pertaining to the other countries are used in the modeling of idiosyncratic risk.

[38] We note that the most significant impact across all assets is that of NFP announcements. Using high frequency data for EUR/USD, Chen, Natividade and Wang [2011] also found this to be one of the events with the greatest impact, second only to FOMC announcement. We note however that the event impact estimation framework here differs in two important ways from that in Chen, Natividade and Wang [2011]: 1) the above variance impact is computed for factor variances rather than asset variances directly, with the events impact for the latter subsequently computed based on the former as explained in the notes to Figure 13; 2) daily (rather than high frequency data) is used.

[39] Event-adjusted returns are equal to the observed returns divided by the square root of the variance multiplier from Step 1 on event days, and left unchanged otherwise.

Figure 13: Event influence on return variance



*Source: Deutsche Bank; We show the US events influence on systematic asset variance. The event impact at asset class level is computed based on the events' impact on factor variance and our factor model as follows: for each asset and each event, we compute the sum of (squared factor beta * factor variance multiplier)/ sum of squared betas. For each event, we plot the average variance multiplier within each asset class.*

4.4 P-Distribution: the Higher Moments

As the reader may suspect, one is unlikely to find a closed-form distribution that accurately reflects all the characteristics of asset returns. But we can get somewhere close using a moment-based approximation. For this, we need two ingredients:

- A moment-based approximation method.
- Tractable, computationally efficient higher moment forecasts.

Regarding the first ingredient, there is no unique recipe, but the *Johnson S.U.* distribution[40] is a convenient option. Not only it is leptokurtic, as are financial returns, but its four parameters[41] can be fit via efficient, quasi-analytical algorithms such as in Tuenter [2001].

A random variable X is said to follow a Johnson SU distribution if:

$$Z = \gamma_{JSU} + \delta_{JSU} \sinh^{-1}\left(\frac{x - \xi_{JSU}}{\lambda_{JSU}}\right),$$

where Z is a standard normal variable, and $\sinh^{-1}$ is the inverse hyperbolic sine function: $\sinh^{-1}(x) = \ln\left(x + \sqrt{x^2 + 1}\right)$. The 4 parameters γ_{JSU}, δ_{JSU}, ξ_{JSU} and λ_{JSU} may be estimated using the moment-matching algorithm described in

[40] See the seminal paper by Johnson [1949] where these distributions were introduced.

[41] While the four parameters do not have an intuitive interpretation, they map explicitly into the first four moments of any target distribution. Here we are mapping the first four moments of *h*-period aggregated GARCH returns.

Although flexible, the main disadvantage of this approach is that a Johnson SU distribution is not guaranteed to exist for any set of mean, variance, skewness and (positive) excess kurtosis. That said, other distributions from the same family can be fit in this case.

Tuenter [2001]. Using the above relation between *Z* and *X*, the *cdf* of *X* can be written as follows:

$$F(x) \approx F_{JSU}(x) = \Phi\left(\gamma_{JSU} + \delta_{JSU} \sinh^{-1}\left(\frac{x-\xi_{JSU}}{\lambda_{JSU}}\right)\right)$$

Differentiating the above, we get the *pdf* of a Johnson SU:

$$f(x) \approx f_{JSU}(x) = \frac{\delta_{JSU}}{\lambda_{JSU}\sqrt{2\pi}} \frac{1}{\sqrt{1+\left(\frac{x-\xi_{JSU}}{\lambda_{JSU}}\right)^2}} \exp\left(-\frac{1}{2}\left(\gamma_{JSU} + \delta_{JSU} \sinh^{-1}\left(\frac{x-\xi_{JSU}}{\lambda_{JSU}}\right)\right)^2\right)$$

where $f_{JSU}(x)$ is the Johnson SU fit to *f*, the *pdf* of *h* -period aggregated returns $R_{t+h,i}$.

One important advantage of our proposed modeling framework (i.e. the factor CCC-MGARCH above) is that the higher moments of $R_{t+h,i}$ - the second ingredient in the approximation framework above – can be obtained in quasi-analytical form for any horizon *h*.[42] The proposed moment-based distribution approximation method together with the computationally efficient method of forecasting higher moments[43] for any horizon complete our P-distribution modeling framework.

4.5 Assessing Accuracy – Mean and Variance Forecasts

Having described the mechanics of our P-distribution, we now evaluate its (historical) accuracy in predicting the future relative to standard benchmarks. We focus on 2 forecasts - the mean and volatility - as these will be most useful for extracting the variance risk premium. The former should help us delta hedge better, while the latter will help decide the direction and size of our VRP positions.

4.5.1 Forecasting Accuracy: the Mean

We start by evaluating the directional accuracy of our forecasts of expected asset returns. We focused on *direction* as opposed to *intensity* as the latter is less relevant for delta hedging. At every evaluation date, we took the sign of the mean of our forecasted return distribution and evaluated whether future cumulative returns were of the same sign (our hit ratio). We evaluated forecast accuracy according to 3 horizons: 1 week, 2 weeks and 1 month. These near-term horizons were chosen so as not to deviate too much from traders' typical risk management constraints.

[42] For certain (univariate) GARCH models, (quasi)-analytical formulae for the higher moments (of the returns and variance processes) are already derived in the literature – see, for example, Alexander et al. (2011). As Simonato (2013) explains, this reduces significantly the computational time. Here, we propose applying his approach to our factor CCC-MGARCH modeling setting.

[43] As specified earlier, quasi-analytical formulae for the higher moments can be derived in our framework. These formulae will be used in our future research involving the P-distribution, but as they are outside the scope of the present paper are skipped here.

We used the same set of 15 representative assets as before, representing equity indices, currencies, commodities and Treasuries. Our forecasts were compared against 3 of our favourite[44] trend-following benchmarks:

- Our naive trend signal, as introduced in Natividade et al [2013] but further modified in Natividade and Anand [2016] to account for asymmetric noise control.
- The Mann-Kendall signal, based on the correlation of sequential returns from different lookback windows. This method, whose details are further explained in Natividade [2012b], has been particularly useful for intraday predictions – see, for instance, Natividade [2013c].
- The Oomen-Sheppard trend signal, based on the relationship between beginning vs end values and peak vs trough values.[45]

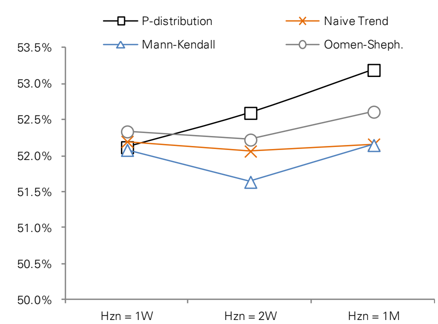
The benchmarks used multiple lookback windows to learn from, and the backtest encompassed 11 years. We addressed the likelihood that spot would mean revert over a certain interval through a non-parametric impulse-response function, which weighted the signal from each lookback window accordingly.

Figure 14 plots, according to each horizon and forecast method, our aggregate hit ratio – the percentage of instances when the direction of a signal is the same as the direction of future asset returns. Two observations stand out:

- All methods seemed to have outperformed a random walk when predicting the direction of future asset returns, regardless of the horizon, over the sample window. That said, the directional accuracy was not far distant from 50%, reiterating the challenges in time series forecasting.
- Our P-distribution forecasts increasingly outperformed the benchmarks as our horizons lengthened. This is unsurprising; most of our signals are of low frequency. While our Sentiment and (some of) our Value signals decay fast,[46] the Trend, Carry and Macro Factor signals target longer horizons. This observation also concurs with our experience trying to explain asset returns at specific frequencies.[47]

These results suggest our forecasts should outperform the futures price and, *to the extent that we can keep our delta positions un-hedged for longer*, that they should outperform other standard methods. This will become more evident in Section 5.7, when we look at the performance of different delta hedging methods across asset classes.

Figure 14: Accuracy (%) of forecasts of the direction of future returns by horizon



We used all 15 markets and 11 years of data. We used hit ratios for accuracy – that is, the % of forecasts of the direction of future returns over a given horizon that turned out to be correct. Source: Deutsche Bank

4.5.2 Forecasting Accuracy: the Volatility

We now turn our attention to the accuracy of our empirical volatility forecasts.

[44] Each looks at trendiness from a different perspective, and the first 2 have been successfully used in our paper trading before. For a larger set of potential benchmarks, see Natividade [2012b].

[45] See Oomen and Sheppard [2014]. We added this method for completeness given that it accounts for trend magnitude in a way that the other benchmarks do not. That said, its estimates suffer from discretization bias as we use very few observations to estimate the signal at any point in time.

[46] It is worth mentioning that weighting our signals unevenly – giving more weight to faster signals – did not improve the forecast accuracy. We would have expected better results on 1-week forecasts had our signals utilised intraday data as well.

[47] In Natividade et al [2014] we ran panel regressions that tried to explain short, medium and long-term contemporary returns in each asset class using a series of market, macro-economic, technical and fundamental variables. We found that the explanatory power rose as we lowered the frequency of one-period returns (from weekly to monthly, then to quarterly, semi-annually and annually).

As with the mean, we evaluate our forecasts under multiple horizons: 1 day, 1 week, 2 weeks, 1 month and 3 months.[48] All 15 markets were included.

We penalized forecasts according to the bias statistic introduced in Connor [2000] – also utilized in Alvarez et al. [2012] and Ward et al. [2016]. This metric measures the standard deviation of returns standardized by the volatility forecasted for the return period.[49] If our forecast is accurate, this standard deviation should equal 1. If we have over-/under-forecasted vol, the computed measure will be less/greater than 1.

We evaluated accuracy in 2 datasets: the long-term data and a subset that isolated crisis periods, as highlighted in Figure 15. In order to capture our forecast accuracy going into these turbulent periods, we started our forecasts one month prior to the dates shown.

Our tests focus on 4 aspects:

1. Parameterization: do parametric models outperform their simpler, non-parametric peers? We tested the GARCH family versus EWMA, classical realized volatility, high frequency volatility and implied vols.
2. Factor modeling: does our factor model outperform its univariate counterparts in predicting asset volatility? We tested both univariate and multivariate versions of our GARCH and EWMA models.
3. Regime modeling: does calibrating our parametric models to exogenous regimes help? We tested our parametric models with and without regime calibration.
4. Event modeling: does separating event-related variance from non-event variance help? We tested both cases.

Figure 16 illustrates all the models used.

We first tested the impact of parameterization. As per Figure 16, we compared both CCC-MGARCH and univariate GARCH – which calibrate for persistence, reaction and level – versus various EWMA estimates, in addition to past realized volatility using both close-to-close and higher frequency data[50] and, ultimately, implied volatility as well. None of these models incorporated regimes or events; these were addressed later.

Figures 17 and 18 show 5 conclusions:

- The GARCH models appear most accurate overall, outperforming their non-parametric peers.
- Within both GARCH and EWMA families, the univariate and factor versions perform similarly.
- The forecast accuracy of all methods worsens as the horizon lengthens.

Figure 15: Market turbulence periods

Period	Event
Sep-/08 - Mar-/09	Lehman + GFC
May-/10 - Sep-/10	EU Sov. Crisis
May-/11 - Jun-/11	EU Sov. Crisis
May-/13 - Jun-/13	Taper Tantrum
Sep-/14 - Mar-/15	Greek Debt Crisis
May-/15 - Aug-/15	China Slowdown
Feb-/16 - Mar-/16	Oil Price Shock
Jun-/16 - Jul-/16	Brexit

Source: Deutsche Bank

Figure 16: Volatility forecasting models

Model	Description	Used in test
RV	Realized volatility (daily data - lookback = forecast window)	1
RV (HF)	Realized volatility (30-min data - lookback = 3 days)	1
Univ EWMAs	7 EWMA models (7 half-lives)	1
Factor EWMAs	7 Multivariate factor EWMA models (7 half-lives)	1
Univ GARCH	GARCH(1,1) model	1, 2
Factor CCC-MGARCH	3 Multivariate factor GARCH(1,1) models *	1, 2, 3, 4
IV	Implied Volatility	1

*Source: Deutsche Bank; * The 3 versions of this modeling framework are: 1) no regimes, no events; 2) no regimes, with events; 3) with regimes, with events. The EWMA half-lives are: 5, 10, 11.2 (corresponds to the Risk Metrics 'daily' lambda of 0.94), 21, 90 and 252 business days.*

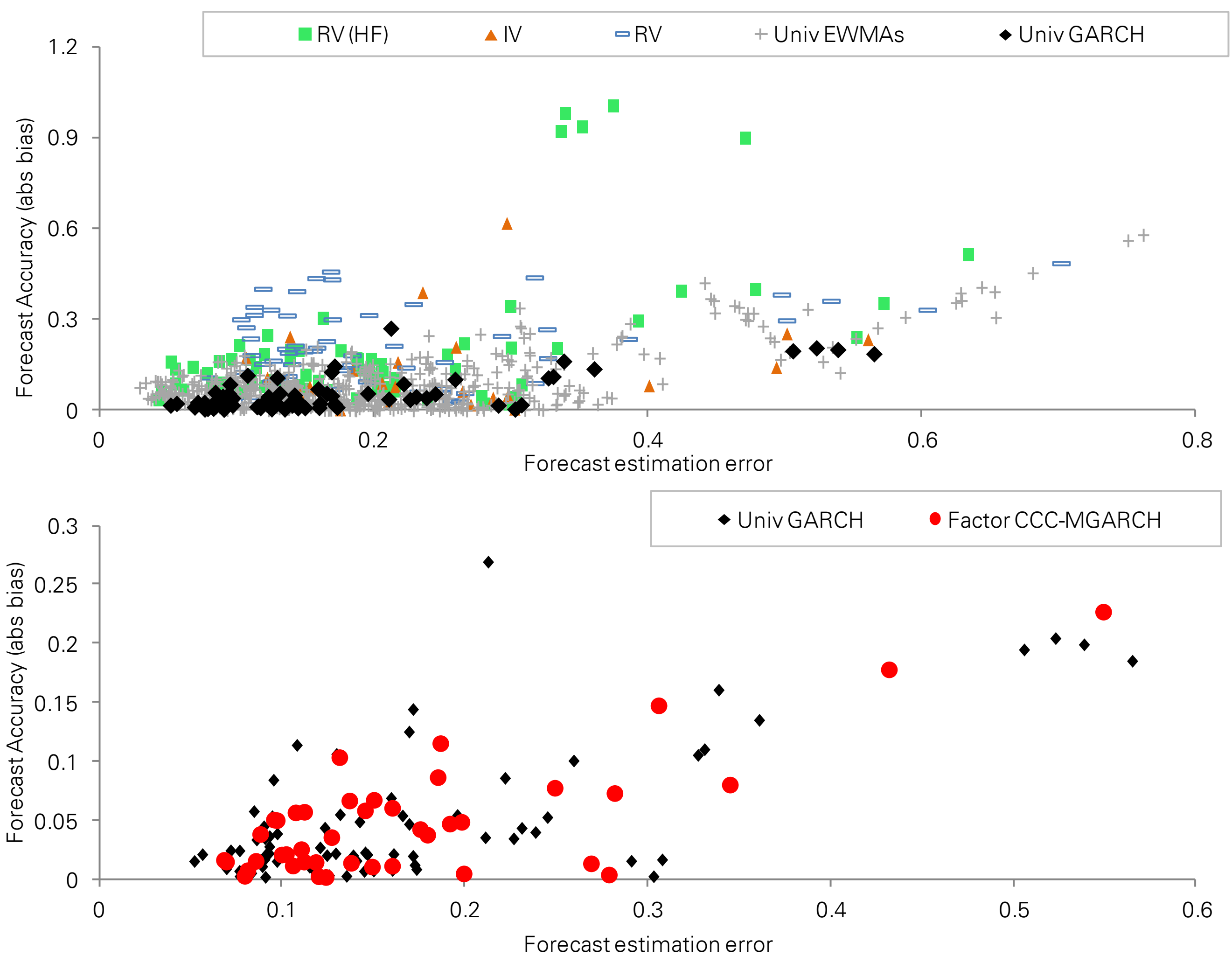
[48] While we used multiple forecast windows, we note that the most relevant forecasts are those up to 1 month. VRP strategies use short-dated options.

[49] As the focus here is on the cross-sectional comparison between models, we used daily observed, overlapping returns in order to improve the sample size at longer horizons, especially for our second data sample. Furthermore, for robustness, we have also run our overall accuracy results on model rankings using non-overlapping weekly returns and our conclusions remained unchanged.

[50] We included "high frequency volatility" estimates as they are popular among market makers and have been advocated in academia – most notably, Ait-Sahalia and Mancini [2008].

- None of the methods accurately predict market shocks; they all under-forecast volatility going into the periods outlined in Figure15. This is a key finding, with implications for signal generation in Section 6.[51]
- That said, implied volatility outperforms the other methods, *when forecasting 3 months ahead*, during those periods. This is likely due to the options market over-forecasting volatility, in general, relative to the other methods.[52] Note also that such conclusion is less clear over shorter horizons; implied volatility does not outperform in 1-month horizons, and we did not include it for shorter horizons.

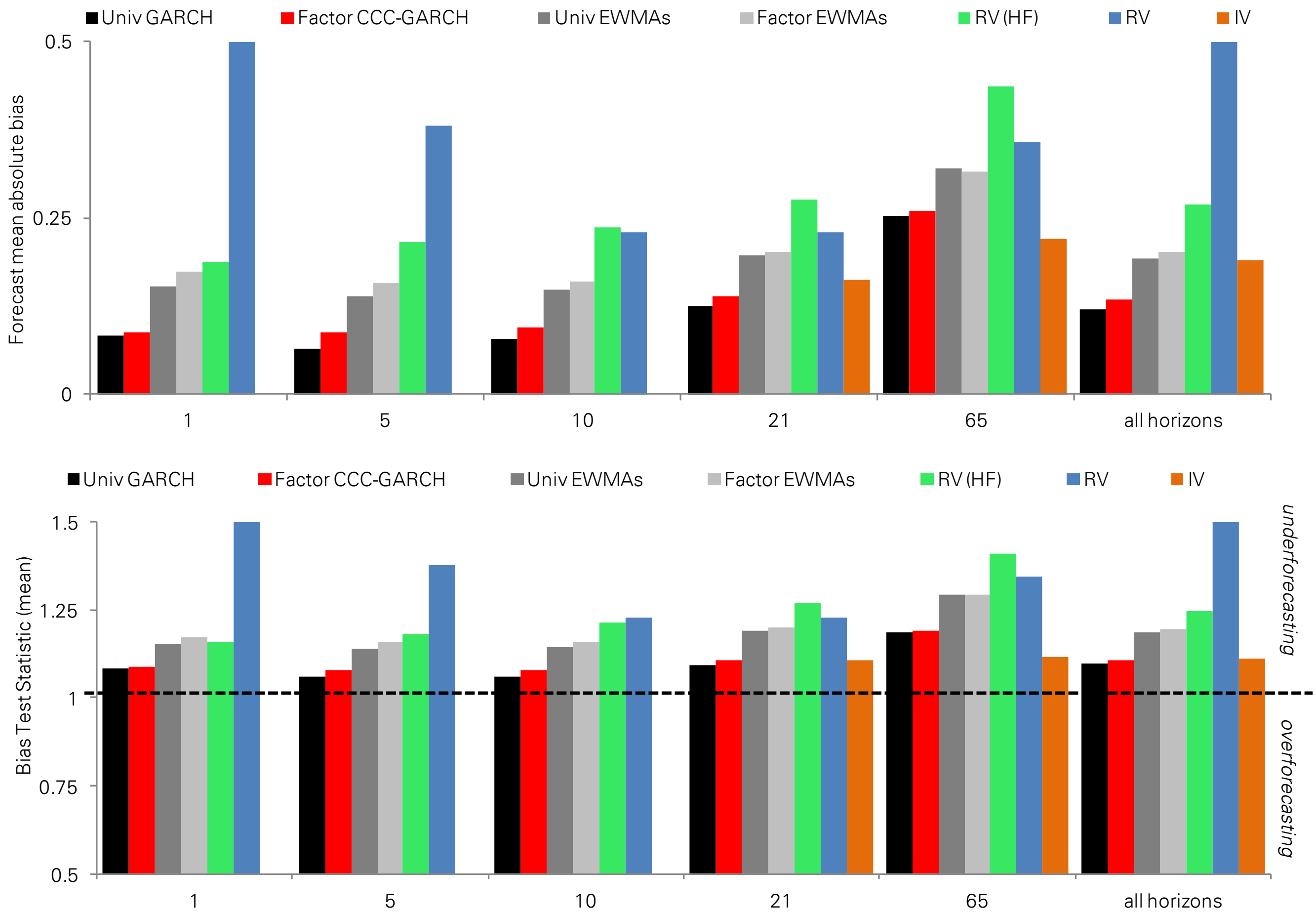
Figure 17: Long-term forecast performance across models (all horizons and underlyings)



Source: Deutsche Bank; Overall data set – Oct 2005 to Aug-2016 – used for computations. Y-axis: we show the 'absolute bias' (i.e. absolute value of the difference between the bias statistic and 1, which corresponds to a perfectly accurate forecast). X: axis: we show the standard deviation of a 252- bd rolling window computation of the bias statistic.

[51] The reader may rightly question what threshold distance from 1 would define statistically significant under-forecasting. We did not delve into that topic, and therefore our statement is generic.
[52] In our view, implied volatilities outperform going into turbulent periods not because the options market is more accurate at predicting shocks, but rather because the time-homogeneous, over-forecasting bias works in its favour in this case.

Figure 18: Forecast accuracy *during crisis periods* according to forecast horizon (days)



The X-axis shows the vol forecast horizons (in business days). Y-axis: top chart: the same absolute bias measure as defined in Figure 17 is now shown averaged across all assets, for each horizon bucket; bottom chart: the value of our bias test statistic (averaged across assets for each horizon bucket) is shown. All computations use the turbulent market periods shown in Figure 15. Source: Deutsche Bank

We opt for the CCC-MGARCH framework – which is both parametric and multivariate. First, we favour its parametric nature as it allows us to approximate the P-distribution more accurately; as mentioned earlier, we need an accurate P-distribution estimate to fully harvest the multiple premias available in options markets. Not only that, but a well-calibrated P-distribution also allows us to simulate asset returns in far more realistic fashion, whose applications go well beyond the trading of options.

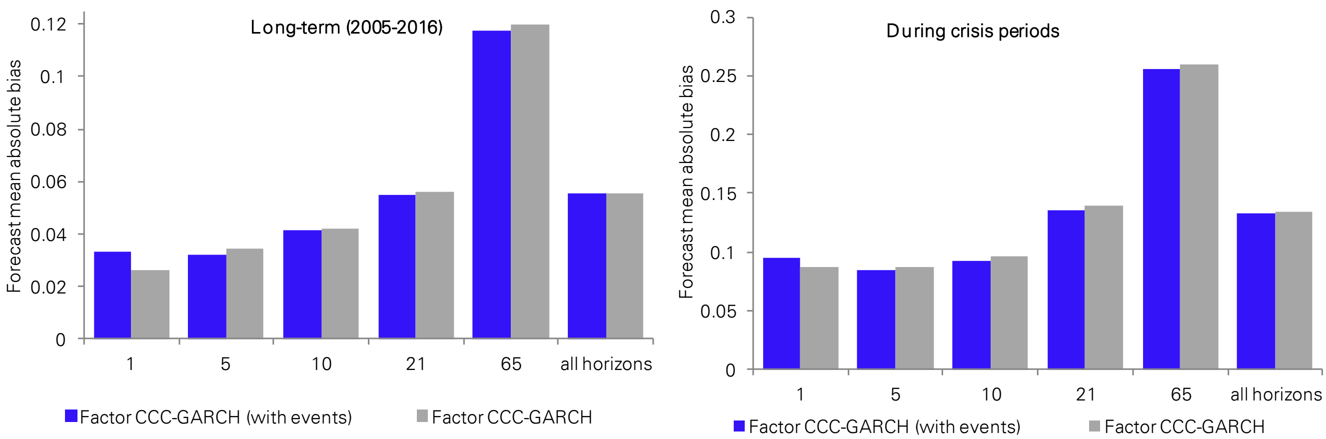
Second, we favour modeling risk through factors because while it looks more complex, portfolio risk estimation is rather simpler. It is no *less* accurate than univariate modeling and has the added benefits of computational efficiency and transparent risk attribution.

We next turn our attention to the modeling of events and regimes, zooming into CCC-MGARCH. Figures 19 and 20 illustrate the following findings:

- Distinguishing between event and non-event returns when predicting future volatility only leads to a trivial improvement, whether in the long run or in turbulent periods. A more granular assessment would also show notable accuracy gains in FX and commodities, as in line with Chen, Natividade and Wang [2011], but these are offset by losses in Equities and Treasuries.

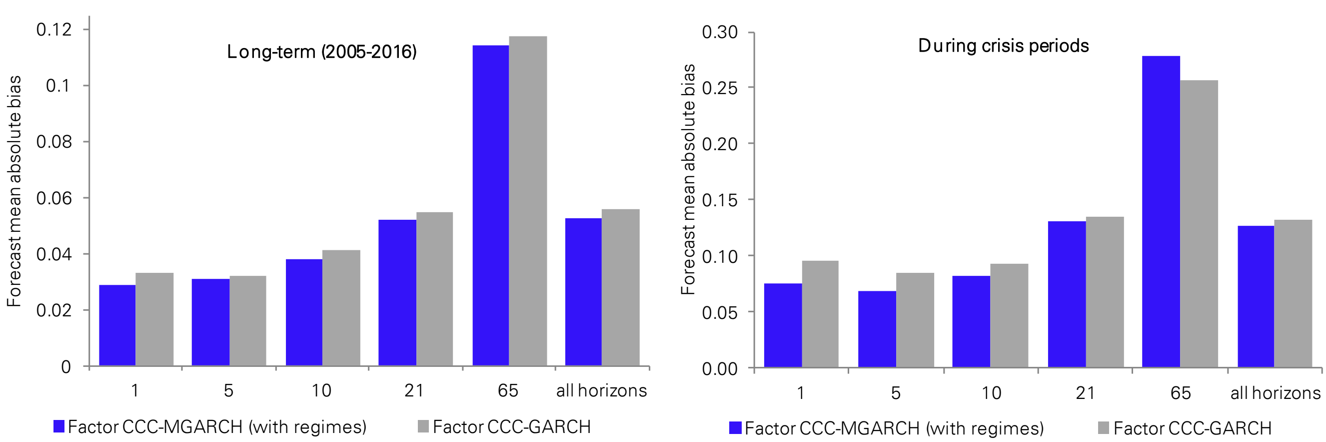
- Regime modeling proved somewhat more fruitful: the CCC-MGARCH model with regimes outperforms its uni-regime counterpart during turbulent markets, for horizons of less than 3 months. A more granular assessment would also show that this is especially true in equity markets.

Figure 19: Event modeling – impact on forecast accuracy according to forecast horizon (days)



Source: Deutsche Bank; Neither model includes regimes.

Figure 20: Regime Modeling – impact on forecast accuracy according to forecast horizon (days)



Source: Deutsche Bank; Note that both models include events.

In light of the results above, we stick to the originally proposed model: CCC-MGARCH, with events and regimes, which will be evaluated for signal generation in Section 6. We now move on to alternative delta hedging, the next step in our quest to improve VRP strategy returns.

5. Alternative Delta Hedging

Alternative delta hedging is, arguably, as ancient as option markets. Market makers will happily deviate from their hedging template if they have an edge in doing so, and if their risk constraints allow. A tailored delta hedging scheme, in essence, improves the positive aspects and reduces the negative aspects of the “textbook” delta hedging approach. For the options seller, for instance, delta hedging is *usually* a loss-making spot-trading strategy: "buy high and sell low". Foresight of future spot conditions could reduce the loss by, for instance, under-hedging[53] if the asset is expected to reverse course. For the volatility buyer, the converse is true. It is true that alternative schemes will increase delta risk, but this side effect should be compensated by the value-add of the new scheme.

The schemes we introduce here utilize our subjective assessment of what the future will look like – specifically, on future spot dynamics.[54] They either focus on the direction of future asset returns, or on how the asset is trending. If correct, using this extra information should give us extra gains.

The trend-based schemes ultimately dictate how much volatility will be captured through delta hedging. They assume that markets are not *fractal*; over the same sample history, high frequency and low frequency returns can be very different.

“True” volatility is a *latent* variable, and therefore what we observe depends on the estimation frequency and method. For instance, if spot appears to move sideways when evaluated in 1-week intervals, but goes through 2-day trends, the annualized 2-day volatility will be higher than the equivalent 1-week volatility. Therefore, if our strategy is to be short options, we should prefer delta hedging weekly rather than daily as in doing so we will capture less volatility and hence incur a lower delta hedging loss.

The trend-based schemes we introduce will, in essence, seek to capture more or less volatility than what is otherwise observed by using close-to-close data points. They do so either by grid-searching optimal delta estimation frequencies or by using moving averages to get different delta quantities.

The direction-based schemes, on the other hand, use a distinctly different approach. Rather than trying to capture more or less volatility, they use

[53] That is, either hedging less often or hedging a level that still leaves her under-hedged.

[54] The reader may question why we did not focus explicitly on volatility – including estimating the delta according to the volatility forecasts introduced in Section 4 and delta hedging daily to it. We see the 2 topics – alternative delta hedging and hedging to realized volatility – as different in nature. The latter is about achieving delta hedging “purity”, while the former is about adding “impurity” to achieve extra gains. We refer the curious reader to Ahmad and Wilmott [2005] and in Derman and Miller [2016], who thoroughly address the topic of delta hedging to “true” volatility. They show that delta hedging to the “true” volatility may provide a terminal P&L that is already known at the start of the trade (the difference between the premium priced using the volatility estimate and the premium currently quoted in the market). The issues, as they partly highlight, are that (1) “true” volatility is often unknown, (2) we cannot delta hedge continuously, and hence suffer from discretization error, and (3) the P&L variance can be quite significant over the life of the trade.

directional forecasts to decide whether to hedge or to run with the current delta exposure. *They allow for more delta risk if – and only if – that risk is likely to pay off*. The delta exposures in this category are managed either through dynamic bands, break-evens and stop-losses, or by using the forecasts from our P-distribution.

Sections 5.1 – 5.5 introduce each scheme, while Sections 5.6 – 5.7 show our backtest results using both simulated and real data. For completeness, we cover both long and short option strategies.

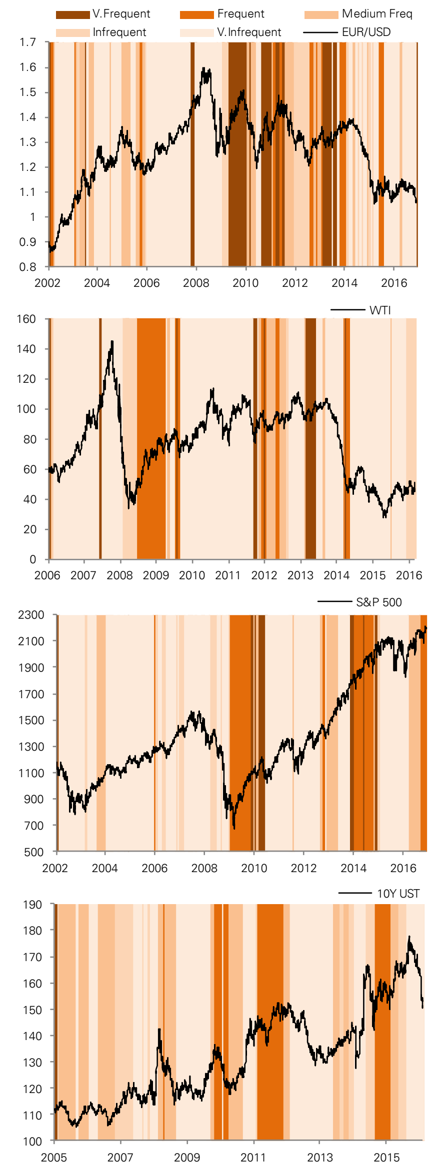
5.1 Grid Search

The grid search is the "brute force" of our alternative delta hedging schemes. Every day, it takes our options basket and calculates what the risk-adjusted returns would have been, up to the day before, from delta hedging that basket at the frequency stipulated by each grid-point.

This method effectively "observes" the underlying market at different frequencies and infers which one would have been best for delta hedging up to time $t-1$. It then decides how to delta hedge at time t.[55] Hedging to frequencies where spot is more volatile should produce better results when we are long options, as we are capturing more volatility. The converse is true when we are short. Gridpoints represent different frequencies, and are divided into 3 domains: *delta*, *time* and *spot*. We describe them as follows:

- Delta frequencies: hedge if the delta of an option has moved by more than a fixed quantity since the last hedge. Delta anchors are popular among market makers, especially when dealing with intraday hedging, as they are akin to trading under a *volume clock*[56]; this frequency domain avoids trading during illiquid periods as both spot and delta stay constant, thereby reducing market impact. The downside, however, is that option deltas can be highly volatile going into expiry – an important issue in practical VRP implementation, as it affects costs and signal-to-noise. The grid points we used are 5-, 10-, 15-, 20- and 25-delta steps.
- Calendar frequencies: delta hedge now if the last hedge occurred after a certain number of days. The results from this domain are easier to interpret, as trendiness and volatility are normally observed as a function of time. Calendar frequencies do not discriminate across liquidity pockets, but these are much less observable in daily data, and daily is the highest trading frequency we assume in this report. The grid points used were 1, 3, 7, 11 and 20 business days.
- Spot frequencies: delta hedge the option only if the underlying asset has moved by more than a fixed quantity since the last hedge. This domain captures the best of the previous two: it captures liquidity pockets and does not suffer from unstable deltas near maturity. That said, it also does not address the path dependency of an option's gamma; this is less of a concern for VRP baskets but more of an issue for individual options. We tested 0.1, 0.3, 0.5, 0.7 and 1 standard deviations (of monthly spot moves) as grid points.

Figure 21: Grid-search preferred delta hedging frequencies



V.Frequent: every 1 day, 3 days, 5-delta or 0.1 standard deviations (in1M spot moves). Frequent: every 7 days, 10-delta or 0.3 standard deviations. Intermediate: 11 days, 15-delta or 0.5 standard deviations. Infrequent: 16 days, 20-delta or 0.7 standard deviations. V.Infrequent: 20 days, 25-delta or 1 standard deviation. As an example, the grid search method recommended hedging every 1 or 3 days, or every 0.1 s.d. changes, during time periods shaded in brown. Source: Deutsche Bank

[55] We used a P&L lookback window of 250 time units – 1 year – in order for the results to be more adaptive to changing market conditions. We used Sharpe ratios as the defining metric. As is standard with grid search methods, we also applied nearest neighbour smoothing both historically and cross-sectionally (within a domain) across grid points.

[56] See Natividade [2013b].

Figure 21 shows the time-varying optimal delta hedging frequency when short a basket of options[57] in four markets, assuming no risk tolerance constraints, overlaid against the futures contract. Darker colours indicate that the grid search favours higher hedging frequencies, while lighter colours show that the method preferred lower frequencies (in delta, calendar and spot domains) at that time.

The results are intuitive, and similar regardless of the asset class. Lower frequency delta hedging has been more optimal because implied volatilities normally overshoot the empirical; the VRP harvester would have been better off just collecting the option premium and rarely trading the (money losing) delta hedging leg. When volatility rises and the asset is trending, however, higher frequency delta hedging performs better as gains from the delta hedge leg offset some of the loss from the short options position.

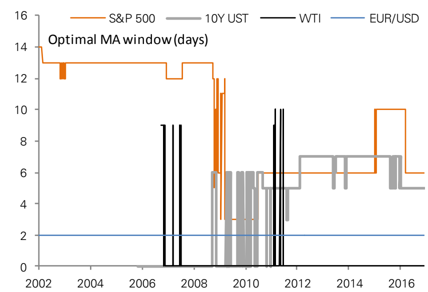
5.2 Moving Averages

The moving average (MA) scheme is similar to the previous in that it also solves for an optimal parameter level, but different in that the parameter in question is the asset price that goes into calculating delta. We hedge the delta daily, but estimate it using a historical moving average price instead of what is currently observed in the market as reference. In essence, this method assumes that it has a better understanding of future conditions than what's "priced in”, and that delta hedging to it will lead to better returns.

The scheme makes certain assumptions about trendiness, volatility[58] and optimal spot references based on the historical performance (up to 1 day prior) of multiple moving average reversal trading strategies[59]. It also picks the best or worst MA as spot reference depending on whether we are short or long gamma:

- If most MA strategy returns are positive, we assume the asset is mean-reverting. If we are short gamma, the delta hedge computation uses the MA level of the *best* MA strategy as spot input. On the other hand, if we are long the option we delta hedge to the current asset price.[60]
- If most MA (reversal) strategy returns are negative, we assume the asset is trending. If we are short the option, we calculate our delta hedge amount

Figure 22: Chosen moving average windows per asset class



Source: Deutsche Bank

[57] We use the baskets introduced in Section 3: 1-month rolling expiry, 12 strikes set 0.5 standard deviations apart from one another.

[58] Changes in the delta, which maps into the P&L of the delta hedge strategy, will be less significant when we use a moving average as input compared to the current spot price.

[59] The MA lookback windows range from 1 day to 1 month; we chose not to use longer windows so as control how much our delta estimate can deviate from the benchmark. The MA strategies are signal-weighted; in other words, the more that spot rises above the moving average, the stronger our short position. Signal weighting is important as delta hedging strategies are gamma-weighted, and gamma intensity relates to the size of recent moves. We use an anchored window to calculate strategy performance, and apply nearest neighbour smoothing as is standard in grid search optimisation. We chose an anchored window so as to capture structural patterns.

[60] If the asset is mean reverting, we expect the moving average to *outperform* current spot at predicting the asset price in the near future. If we are short a call, and spot has been rising, we will *buy less* spot in the delta hedge than the original because we will use a lower level of spot as input. This decision to under-hedge will be correct if spot is indeed mean reverting and therefore subsequently drops, as we will lose less from the previous trade. If we are long a call, we will delta hedge to the current asset price. If we instead delta hedge to a moving average, we would end up *selling less* than the benchmark trade and not capitalise as much when the asset price eventually drops, which is sub-optimal. *We choose the best performing MA as we assume it is the one that spot is more likely to revert to.*

- using the current asset price as input. If we are long the option, we delta hedge to the current MA level of the *worst* performing MA strategy.[61]

Figure 22 plots the moving average windows chosen to estimate the level of spot (or futures) used in our delta hedge calculator, assuming that we are short options. If the chosen MA window was zero, it meant the asset was trending in the short-run. In that case, we assumed that current spot would outperform a historical average when predicting future spot, and thus chose not to under-hedge our deltas. The opposite applied if the asset was mean-reverting, and the length of the optimal moving average dictated how under-hedged we were.

A key highlight from Figure 22 is the stability of our results for the S&P 500, an asset known for its short-term mean-reverting properties. In this case, delta hedging to a 5-10 day moving average is generally preferred.

Currencies, on the other hand, tend to mean-revert under shorter windows, which is also reflected in our choice for the 2-day lookback. As currency markets often trend under shorter horizons than equities[62], it is unsurprising that the algorithm opts for smaller MA lookback windows – we should under-hedge by less.

5.3 Break Outs

The break-out scheme is comprised of 2 variables: the options break-even and a rolling stop-loss on spot. Well-suited for a market maker, it sits in between the two smart delta hedging categories described above.

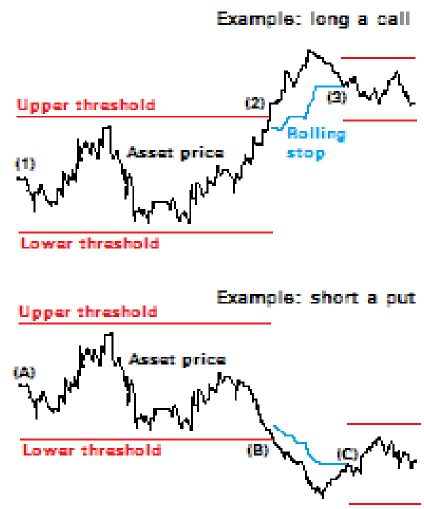
This scheme resembles trend following algorithms in that it uses a break-out filter and a *stop* mechanism, be it *stop-loss* or *stop-gain*, just as classical trend following systems do.[63] The break-out bands, in this case, are equal to spot plus and minus the price of an ATMF option whose expiry equals that of the option we are delta-hedging. We interpret these bands as *the option market's best estimate of the current break-out ranges in spot.*

It is also similar to the direction schemes we describe later because once the asset breaks outside the aforementioned range, we assume that it will continue moving in that direction, and therefore increase or decrease the delta hedging frequency depending on our option position.

The scheme works as follows:

1. At inception of the option trade, or when the rolling stop has been triggered, we calculate the spot range: spot +/- the premium of an ATMF option that expires the same day as the option we are trading.

Figure 23: Delta hedge procedure under the break-out method



(1) Establish the break-out thresholds as asset price +/- premium; delta hedge daily. (2) Spot has broken to the topside: stop delta-hedging, initiate rolling stop. (3) The stop-profit got triggered: establish the new break-out thresholds and delta hedge daily again. (A) Establish the break-out thresholds as asset price +/- premium; do not delta hedge. (B) The asset price has broken to the downside: start delta hedging daily, initiate rolling stop. (C) The rolling stop has been triggered: establish the new thresholds and stop delta hedging.Source: Deutsche Bank

[61] If the asset is trending, we expect the moving average to *underperform* the current spot level at predicting the asset price in the near future. In reverse image to the Footnote above, if we are short a call and delta hedge to the moving average we will end up buying less spot than the original and therefore not capitalise enough as the asset continues to rise. In this case, therefore, we use the current asset price as input to calculate the delta hedge. On the other hand, if we are long a call, we would prefer being under-hedged as spot keeps rising, and delta hedging to the moving average gives us that. In other words, we sell less spot than the benchmark delta hedge strategy suggests. *We choose the worst-performing moving average as it is the one that spot is the least likely to revert to.*

[62] These results are consistent with our work on impulse response functions shown in Natividade et al [2014b].

[63] See, for instance, Clenow [2013].

2. While the spot level is inside the range, if we are long the option, we delta hedge daily so that the gain from our long gamma exposure provides relief against the loss from our short theta exposure. In other words, we "gamma scalp". If we are short the option, however, we do not delta hedge and collect the returns associated with time decay instead.
3. If spot breaks to the topside of the range, we initiate a rolling stop trigger set at 0.3 standard deviations below spot and roll it daily.[64] The delta hedge varies according to the option position:
   - If we are long a call or a put: stop delta hedging[65] and "let the delta run".
   - If we are short a call or a put: cover the delta exposure immediately and start delta hedging daily.
4. If spot breaks to the downside of the range instead, we also start a rolling stop level. The delta hedge varies:
   - If we are long a call or a put: stop delta hedging and "let the delta run".
   - If we are short a call or a put: cover the delta exposure immediately and start delta hedging daily.
5. If spot retraces and hits the rolling stop, we cover the delta exposure immediately and re-start the process above.

Figure 23 illustrates the process when applied to long call and short put positions, under opposite market scenarios.

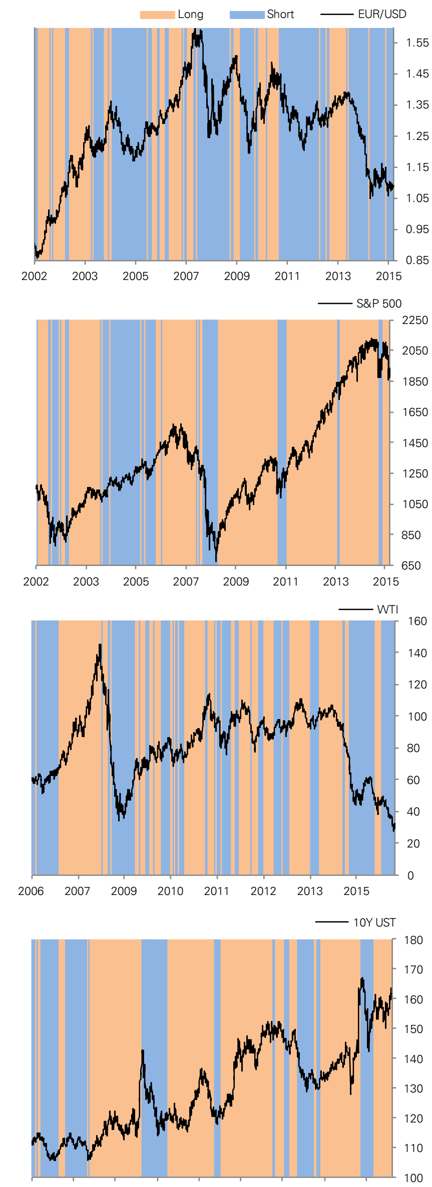
5.4 Expected Returns

The Expected Returns (ER) scheme explicitly incorporates our return direction forecasts into the delta hedging decision, more so than any other scheme that we are introducing. The idea is simple: *if the delta hedge trade we need to do today is in line with our spot view, we delta hedge; otherwise, we do not.* In other words, we will stay under-hedged or over-hedged if that suits our spot views.

The better our forecasts can predict the direction of future spot returns, the more that this delta hedging approach should outperform the others. That said, it also causes our VRP strategy to deviate from the benchmark, which – in extreme form – may lead to a significant breakdown in correlations with pure VRP returns.[66]

Our return forecasts come from the P-distribution introduced in Section 4. As described earlier, the forecasts combine all the individual signals from our delta-one portfolios – Trend Following, Carry, Value, Sentiment and Macroeconomic Factor investing. We are primarily interested in the direction of

Figure 24: Signals from the Expected Returns scheme: long vs short the underlying asset



Source: Deutsche Bank

[64] The length used to calculate the standard deviation is equal to the number of days between the last delta hedge date and today. The stop adjusts itself at every spot move but cannot become *less* stringent than it was before. The 0.3 s.d. level was not optimised to this exercise; it was similar to a choice we used before. See Natividade [2013c] for more details.

[65] In cases where we are long the option and stop delta hedging, we do not remove the previous hedge. We simply stay under-hedged from this point onwards.

[66] In this case, we could end up capturing probability risk premium instead of the volatility risk premium.

these returns, as opposed to the magnitude, as the former suits delta hedging best.

Figure 24 plots flagship assets in FX, equities, commodities and Treasuries against a background whose colours reflect the output from our spot models – bullish (orange) or bearish (blue). One can see that the directional views from our ER method do not flip often, a potential downside given that delta hedging normally classifies as short-term trading.

5.5 Whalley & Wilmott

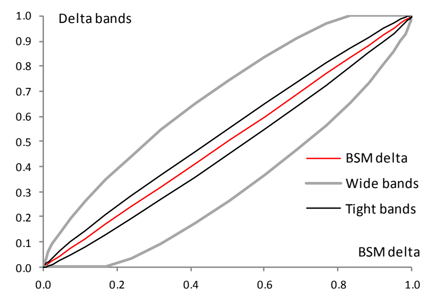
Finally, we introduce a variation of the Whalley & Wilmott (WW) scheme. This method has its origins in academia[67], having been first introduced in Whalley & Wilmott [1997]. It is a transaction cost model that defines a *no-transaction* region for delta hedging. The costlier it is to trade the underlying, the wider the band and therefore the lower the trading frequency. In our context, we modify the model such that the delta bands reflect our views on the underlying asset instead of trading costs. The bands are defined as follows:

$$Bands = \Delta_{BSM} \pm \left(\frac{3}{2} \frac{\exp\{-r(T-t)\}\lambda S\Gamma^2}{\gamma} \right)^{1/3}$$

where λ is a function of the transaction cost[68], r is the risk-free interest rate, t, T are time references, S is spot, Δ_{BSM} is the Black-Scholes-Merton delta and Γ is the BSM gamma.

The variable of most interest to us is γ, a risk aversion parameter[69] that defines the convexity of the bands. Instead of setting it as a function of our cost-related utility preferences, as done originally, we define it according to the spot views coming from our P-distribution. In essence, this method is a *constrained* version of the ER method – we allow for *some* freedom in our delta hedge decisions but ultimately also apply exposure thresholds. For instance, if today's delta hedging trade is in line with our view on future asset returns, we apply very tight delta hedging bands. Otherwise, we set them wide.

Figure 25: Delta hedging bands – Whalley & Wilmott method



Source: Deutsche Bank

[67] To our knowledge, *smart* delta hedging is not a popular theme in academic circles. Delta hedging discretization attracted a good deal of attention in previous decades, but the techniques introduced sought to reduce trading costs as opposed to capture directional views or the non-fractality of asset returns.

The topic was initially addressed in Leland [1985], where the author revised the option's implied volatility to account for the costs of delta trading. Hodges and Neuberger (HN) [1989] followed through with an optimal but computationally expensive band-like approach. Whalley and Wilmott [1997] refined the method with an analytical formula, and Zakamouline [2006a, 2006b, 2007] further expanded on it using a double asymptotic method that more closely resembled the optimal HN approach. We prototyped the 2 latter methods; while the Zakamouline scheme seemed most appropriate for the cost problem, it added little extra to WW while also costing more computational time. Sinclair [2008] provides a thorough review of the topic.

The main references we found on *smart* delta hedging are Chen [2010] and Chen et al [2011], where the author identified a jump-based delta hedging rule applicable to options on S&P 500 futures. Chen argues that as the S&P 500 is likely to revert after large jumps, one should suspend rebalancing the hedge portfolio after large jumps and only rebalance the next day. Finally, Sepp [2013] implemented a grid search method similar to what we introduced in Section 5.1, but with the goal of reducing delta hedging costs as opposed to alpha generation.

[68] $tc = \lambda \mid N \mid S$

[69] Refer to Whalley and Wilmott [1997], Sinclair [2008], Chen [2010] and Chen et al [2011]for details.

Figure 25 illustrates the method in more detail, where the delta bands are plotted against the original Black & Scholes delta.

5.6 Backtest Results – Simulated Data

Having introduced our alternative delta hedging schemes, we now evaluate individual performances on both simulated and real market data. We are ultimately interested in the latter, but the former will be the backbone of our understanding of how different methods perform under various conditions.

We focused on 2 types of simulated data – *trending* and *mean-reverting*. We simulated slightly over 4,000 datapoints in each[70], using AR(1) return generation processes as in Kaminski and Lo [2007]:

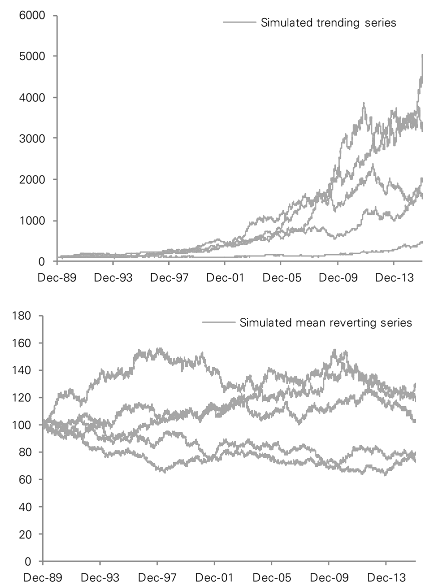
$$r_t = \mu + \rho(r_{t-1} - \mu) + \varepsilon_t$$

where r_t are daily asset returns, $\varepsilon_t \sim \mathrm{N}(0, \sigma_\varepsilon^2)$ is a noise parameter and $\rho \in (-1,1)$ reflects the mean reversion rate. Our trending price series are set such that annualized asset returns and volatility both average 10%, and $\rho = 0.5$ to indicate strong trendiness. Our mean-reverting series also have 10% annualized volatility, but average returns are 0% and $\rho = -0.5$ to indicate strong reversal.[71] Figure 26 shows 10 runs of our simulated trend and mean-reverting series.

As for backtesting our options strategies:

[1] We used the same basket of options introduced at the start of Section 3: twelve 1-month options, whose strikes are set 0.5 standard deviations apart from one another. We assumed a flat volatility surface such that all options were priced at 10% implied volatility. Other operational details were standard for simulation: zero interest rates, zero transaction costs, individual delta hedging for each contract, and roll at expiry.

Figure 26: Simulated trending and mean reverting price series – 5 paths each



Source: Deutsche Bank

[70] If each time unit represents a day, this would be equivalent to 16 years of data.

[71] To make it easier for the reader: $\sigma_\varepsilon = \sigma_A \sqrt{1+\rho^2}$, where $\sigma_A = 0.1$. Further, $\mu = \mu_A / 252$, where $\mu_A = 0.1$ in the trending series and $\mu_A = 0$ in the mean reverting series.

[2] We evaluated 4 strategies: [a] long the options basket in a trending market, [b] short the basket in a trending market, [c] long the basket in a mean-reverting market, and [d] short the basket in a mean-reverting market. Figure 27 provides a sketch of that.

[3] We also imposed *delta risk* thresholds so as to respect the typical constraints of a market making desk. As such, the strategy automatically delta hedges the moment that our exposure exceeds the threshold, whether or not it is in line with the candidate method. The thresholds used were 5-, 10-, 20- and 100-delta, where the latter represents the unconstrained version. This resulted in 80 backtests: 4 strategies, 4 tolerance thresholds and 5 candidate hedging methods.

[4] All results were ultimately compared to the benchmark: the same options strategy but delta hedged daily instead.

We assessed 2 P&L characteristics in each backtest. Our *utility* metric was the marginal Sharpe ratio gain versus the benchmark; other metrics, such as drawdown-related performance, did not suit this controlled environment. Our chosen *risk* metric was the correlation between strategy and benchmark; we wanted to penalize strategies that captured less of the VRP phenomena but not penalize strategies that are simply more volatile. The benchmark strategy is to delta hedge daily.

Figure 28 illustrates the backtest results on simulated data. Each sub-figure aggregates the same data but via different buckets: returns by strategy and market environment ([2]), by delta tolerance threshold ([3]), and by delta hedge scheme. The Y- and X-axis display the utility and risk metrics defined above, except for we display “risk” as 1 minus correlation for better visualization.

Delta hedging strategies closer to the top left should be preferred. Our key findings are that:

- *Alternative delta hedging* outperformed the benchmark in 55% of the instances and in 3 out of the 4 strategy environments described in [2]. The benchmark strategy – delta hedging daily – ranks last among schemes in risk-adjusted returns.
- In the large majority of instances, our backtest returns were heavily correlated to the benchmark.[72] In other words, all methods and most tolerance thresholds still allowed us to capture VRP returns just as we did when delta hedging daily. A tolerance threshold of 20-delta mildly outperformed others from a utility-versus-risk perspective. But as expected, the tracking error rose as we loosened our delta tolerance thresholds.
- Our alternative delta hedging schemes performed best when the asset was mean reverting and we were short options ([d]), followed by when the asset was trending and we were long the basket ([a]). In both instances, the methods *under-hedged*, thereby “letting the delta run”.
- The only bad environment for our methods was [c] – long the basket when the asset is mean-reverting. This is no surprise; it is optimal in this environment to delta hedge as often as possible, which means delta hedging daily. The benchmark, which does only that, is therefore the best “gamma scalping” method.

Figure 27: Backtest scenarios – simulated data

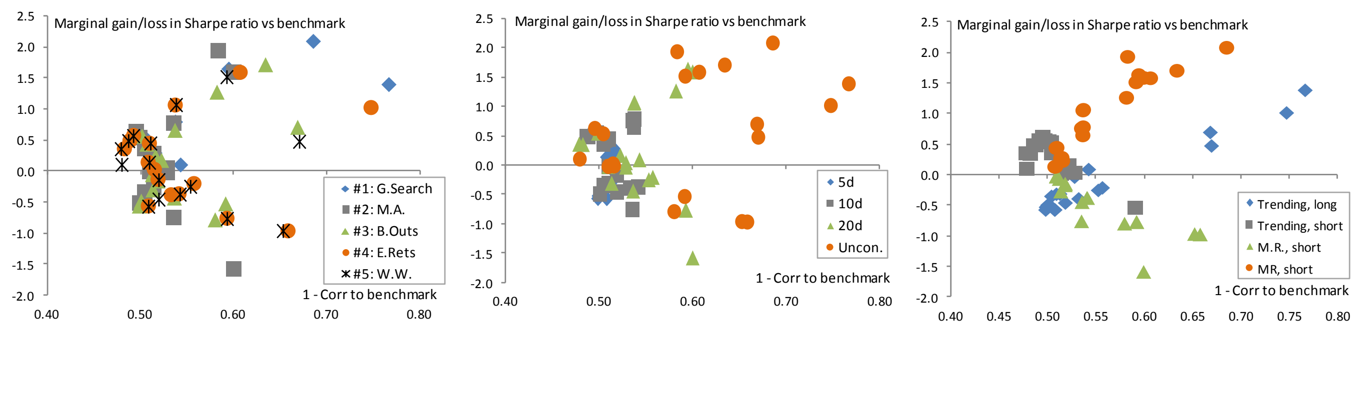
	Trending asset	Mean-reverting
Long gamma	[a]	[c]
Short gamma	[b]	[d]

Source: Deutsche Bank

[72] The correlation of daily log changes, using 4,049 units, was above +0.4 in most instances.

- No scheme subsumed the others, as seen by the lack of clustering in Figure 28. Grid Search (5.1) ranked the highest, which means that the optimal hedging frequencies of the past also outperformed in the future. Again, no surprise – we used a controlled backtest environment with stationary returns, subdued noise and no regime shifts.

Figure 28: Backtested results by hedging scheme (left), delta tolerance (middle) and market environment (right)



Source: Deutsche Bank

5.7 Backtest Results – Real Data

Now that we have an idea of how the results *should* look like, it is time to evaluate how they actually look. Using up to 15 years of data, we applied all delta hedging schemes to each of the 15 markets – in other words, the cross-asset pool described earlier. The backtest setup was similar to the previous in that we used the same (1-month) options basket, the same rolling frequency, the same evaluation criteria and the same benchmark – the returns from delta hedging the basket daily.[73] As before, the delta hedge applied separately to 2 combinations: long the options basket systematically, and short the options basket systematically. As before, we did not add costs – this will be left to Section 6.

Our results are presented in Figures 29 and 30. Figure 29 shows the central finding from our (near) 600 backtest iterations: *on average, the Expected Returns (5.4) and Whalley & Wilmott (5.5) schemes outperformed the others and the benchmark, whether we were long or short options*. In addition:

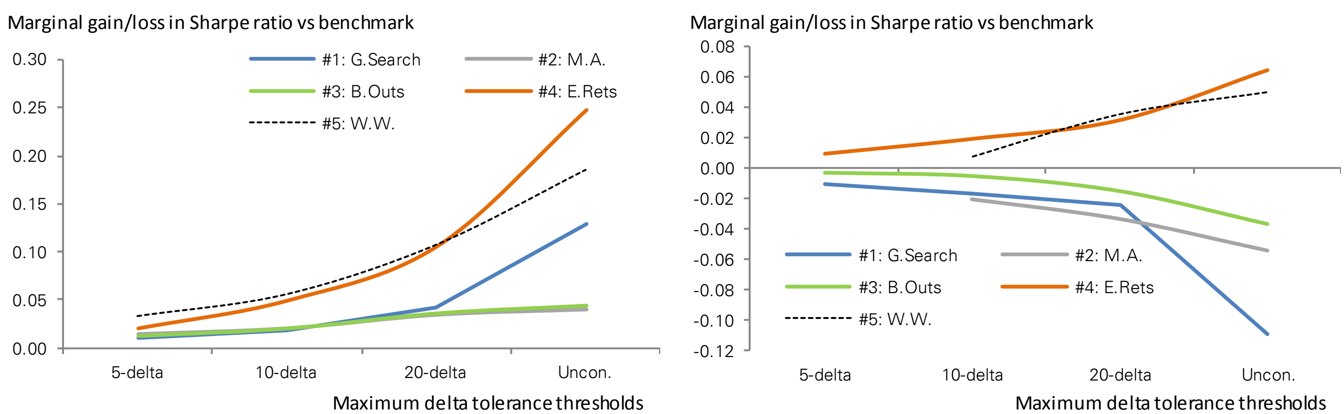
- All schemes outperformed the benchmark when short gamma. We find that intuitive as it resembles the results from scenario [d] in Figure 27: short options in markets whose returns are mean-reverting – in this case, over the short run. Here, the delta hedge strategy buys more in rising markets and sells more in falling markets. The less we trade it, the more under-hedged we are, and the less we lose relative to the benchmark. On average, all schemes under-hedge the delta relative to the benchmark.
- Most schemes underperformed the benchmark when long gamma. This result is similar to that of simulated scenario [c] – long options in (short-

[73] As before, we only look at the delta hedge leg of the returns, thus ignoring the option mark-to-market leg. The latter is the same for all strategies.

term) mean-reverting markets. None of the schemes captures this reversal pattern as aggressively as the benchmark, which increases the size of its short (long) position every day that the asset rises (falls).

- The amount of outperformance and underperformance grew as the tolerance threshold for delta exposure grew. This is in line with the arguments above, and with our findings using simulated data.
- Contrary to the others, the Expected Returns (5.4) and Whalley & Wilmott (5.5) schemes excelled when long and short options. Further, the outperformance grew as the tolerance thresholds grew. In our view, this is because the schemes are based on the directional forecasts of our P-distributions, and therefore account for a much broader range of asset returns than short-term price action alone. The results indicate that these schemes can add value even when daily hedging already looks ideal. Finally, the fact that these schemes further outperformed under looser delta thresholds is in line with Figure 14, which shows that our return forecasts improve over longer horizons.
- Unlike with simulated data, the grid search (5.1) scheme underperformed in the new backtest. It shows that naïve calibration is often not enough when dealing with real data; persistent patterns in asset returns can be clouded by regime shifts and intermittent drivers.

Figure 29: Aggregate backtest results (short – left chart, long – right chart) according to delta hedging scheme, expressed relative to the benchmark



Source: Deutsche Bank

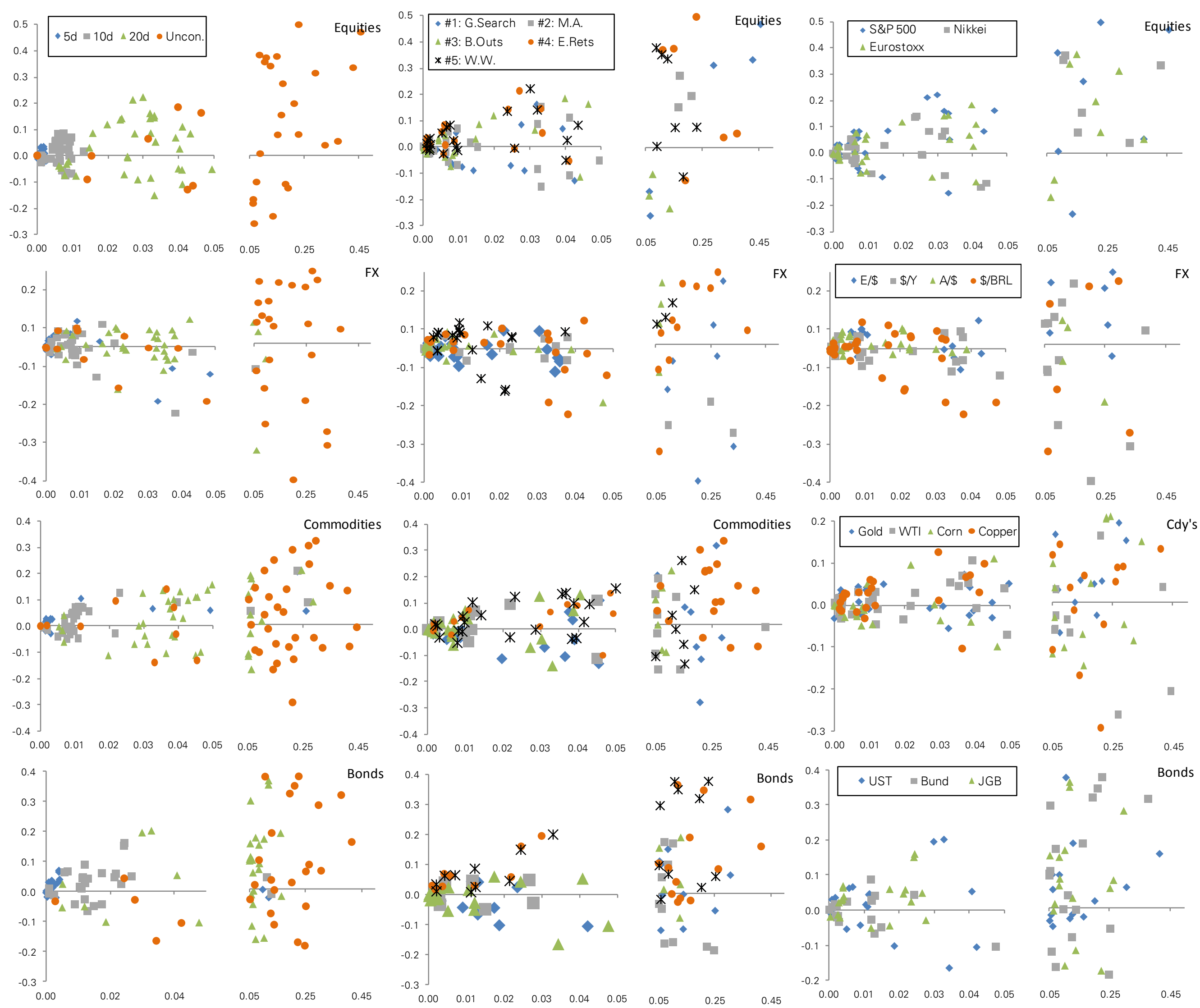
Figure 30 shows additional results. As before, the Y-axis displays the incremental Sharpe ratio from using the aforementioned schemes versus the benchmark strategy. The X-axis shows how the returns deviate from the pure VRP benchmark, as we subtract the correlation between scheme and benchmark returns from one. The first column of charts shows each backtest labeled according to delta thresholds, while the second divides our results according to hedging scheme and the third partitions them according to underlying markets. The rows represent individual asset classes.

The following observations stand out:

- The more we loosened our delta tolerance thresholds, the more that the tracking error grew irrespective of the hedging scheme. This is the most discernible pattern, as shown by the charts on the left and in line with the backtests on simulated data. However, *we are still capturing the VRP – all correlations to the benchmark*[74] *remained above +0.5*.

- Schemes from Sections 5.4 and 5.5 are located, on average, above others and less often below the X-axis. The charts in the middle column show that.
- The charts on the right show no obvious pattern. This is good; we do not want our results to be biased by a small number of assets. It also shows that we are capturing a feature of the VRP premia as a whole, and not tailored to a specific market.

Figure 30: Aggregate backtest results according to delta hedging scheme, expressed relative to the benchmark. Y-axis: marginal gain/loss in Sharpe ration vs benchmark. X-axis: 1 – correlation to the benchmark.



Correlation of daily returns, 8-15 years of data depending on the market. Source: Deutsche Bank

[74] Daily returns, 8 to 15 years of data depending on the underlying market.

6. Building the Strategy

So far we have shown that there is value in forecasting asset volatility and in delta hedging differently. We now evaluate whether these can be used to build a better volatility risk premium signal and, ultimately, a better VRP strategy.

Our tests covered the same 15 markets as before, and 1-month options. But as this section deals with implementation, we modified the variance swap-replicating options basket to be comprised of 10- and 25-delta puts and calls, and ATMF puts and calls, all rolled at maturity. The weight of each option was inversely related to strike levels, as per standard variance swap replication. We also added transaction costs; we used our internal database to estimate volatility bid-ask spreads for the key strikes, and applied linear interpolation to cover other points in the bid-ask surface. As was the case before, our benchmark is a strategy that sells the same basket of options every month, delta hedged daily.

This section covers signal estimation and timing, in addition to implementation aspects of our alternative delta hedging (ADH) strategy and how it interacts with timing. We finalise by presenting our results on selling option baskets on 3 equity indices: the S&P 500, Eurostoxx 50 and the Nikkei 225.

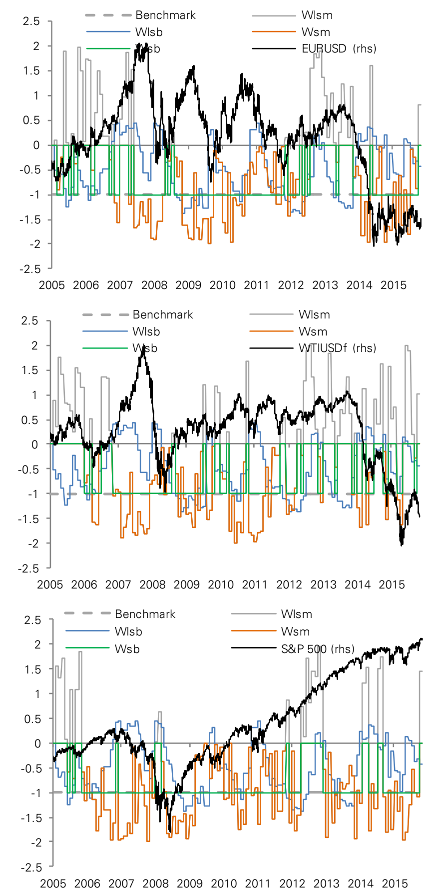
6.1 Fine tuning the VRP signal

We first focus on the volatility forecasts from Section 4. As CCC-MGARCH outperforms in forecasting future volatility, it may also produce better volatility trading decisions than the benchmark – which always sells a constant amount. Here we assumed daily delta hedging but changed the size of our short (and long) options positions to account for the forecasts. The weighting schemes we tested are a function of the spread between current implied[75] volatility and our forecasts of future realized volatility.[76] They focused on direction (long or short, short-only) and magnitude (binary weights, weight by signal intensity).

The weights were built as follows:

1. We took the absolute spread $|s| = \left| \frac{\sigma_I - \sigma_{4C}}{\sigma_I} \right|$ and calculated the current percentile relative to the past 1 year[77] – our signal p.
2. We turned the signal into positions according to direction and magnitude:
   - Long-short, binary-weighted: $W_{lsb} = -sign(s)$.
   - Long-short, magnitude-weighted: $W_{lsm} = p \times W_{lsb} \times 2$.[78]

Figure 31: Signal-weighting schemes



Source: Deutsche Bank

[75] We used variance swap strikes where available, typically sourced from the CBOE. Where unavailable, we used a weighted average of 1-month implied volatilities across the smile so as to proxy the variance swap strike.
[76] It is worth noting that we also backtested variations where the original signal was based either on our volatility forecast alone, or implied volatilities alone. While these schemes were more exposed to factor momentum, the results were not significantly different in how they compared versus the benchmark.
[77] We use a 1-year lookback window to make the signal more adaptive.

- Short-only, binary-weighted: $if\ sign(s) \leq 0, then W_{sb} = 0, otherwise W_{sb} = -1.$
- Short-only,magnitude-weighted: $W_{sm} = p \times 2 \times W_{sb}$

Figure 31 plots our weighting schemes over time in 3 markets – EUR/USD, S&P 500 and WTI – overlaid against asset returns. All 3 assets have typically fallen faster than they rose in the past, which results in most long volatility positions coinciding with bearish spot markets.

That is not to say, however, that the long options positions paid off. This is best seen in Figure 32, which ranks the backtest results from different weighting schemes according to their Sharpe ratio and shortfall-over-returns. The benchmark strategy (short-only, delta-hedged daily) ranks first. The W_{sb} (short-only, binary weighted) scheme, which uses the least amount of information from our signal p, ranks second.

In other words, neither the long volatility positions nor the long volatility *predictions* translated into overall outperformance. As alluded to in Section 4.5.2, our forecast method – and all other benchmarks, including implied volatility – are not entirely accurate at predicting market shocks. Whether magnitude-weighted or binary-weighted, the long positions added more noise than value.

The short-only weighting schemes outperformed the long-short, but still trailed the benchmark. This is because if the volatility forecasts went above the market implied, the exposure went to zero. These flat positions often translated into missed opportunities to sell volatility and capture the VRP.

Finally, the new weighting schemes led to a reduction in return correlations to the benchmark strategy.[79] By using our forecasts to reduce losses or turn them into gains, we ended up less exposed to the risk premium we aimed to capture in the first place. VRP strategies are not *alpha* strategies; in our view, they should not be designed to produce positive returns in every market environment.

These results highlight a challenge often faced by quantitative researchers: *better statistical results may not lead to better strategy returns*. CCC-MGARCH may outperform its peers at forecasting volatility, but it still suffers from the same general shortcomings that make it underperform, *on average*, the naïve benchmark strategy.

We will use this forecast method elsewhere – it is a key input to our P-distribution, and as such will be used for deriving probability risk premium signals and any other use of our P-distribution technology. For volatility risk premium, we opt instead for the original signal and therefore will hold a constant short position on the options basket.

Figure 32: Performance rankings versus the benchmark

	Benchmark	LS M. (Wlsi)	S M (Wsi)	LS B. (Wlsb)	S B (Wsb)
Sharpe ratio					
S&P	3	5	4	1	2
SX5E	2	4	1	5	3
Nikkei	1	4	3	5	2
Bov	2	5	4	3	1
E/$	1	5	3	4	2
$/Y	1	5	3	4	2
A/$	4	3	5	1	2
$/BRL	1	4	2	5	3
Gold	1	5	4	3	2
WTI	1	5	3	4	2
Corn	4	2	1	5	3
Copper	5	1	2	4	3
UST	1	5	2	4	3
Bund	3	2	1	5	4
JGB	1	5	4	3	2
Rank of Avg	1	5	3	4	2
Shortfall over annualised returns					
S&P	3	5	4	2	1
SX5E	1	4	2	5	3
Nikkei	1	4	3	5	2
Bov	2	5	4	3	1
E/$	1	5	3	4	2
$/Y	1	5	3	4	2
A/$	4	3	5	1	2
$/BRL	2	4	1	5	3
Gold	1	5	4	3	2
WTI	1	5	3	4	2
Corn	5	2	1	4	3
Copper	5	1	2	4	3
UST	1	2	5	3	4
Bund	3	2	1	5	4
JGB	1	4	5	3	2
Rank of Avg	1	5	3	4	2

Sharpe ratios are ranked in descending order, while Shortfall-over-returns are ranked in ascending order.LS M: long-short, magnitude weighted. S M: short-only, magnitude-weighted. LS B: long-short, binary weighted. S B: short-only, binary weighted. Backtests since 2005, where available. Source: Deutsche Bank

[78] We multiply the signal scores by 2 so that the absolute average historical weight is 1; this allows for better comparison against the benchmark.

[79] The 11-year correlations of daily returns versus the benchmark VRP strategy were 0.43 for the long-short weighting schemes and 0.78 for the short-only weighting schemes.

6.2 Adding alternative delta hedging

Our alternative delta hedging methods are next. The results from Sections 5.7 are encouraging, but must be understood in more detail. Given its outperformance, we opt for hedging scheme from Section 5.4 (Expected Returns) but also cap our absolute delta exposures to a maximum of 20 delta so as to reduce deviations from the benchmark. Whenever the threshold is hit, we automatically hedge the delta back to zero.

The top chart in Figure 33 shows that a VRP strategy that delta hedges according to the scheme above outperforms the benchmark in almost every market. This had been alluded to in Section 5.7, and is attributed to the predictive power of our P-distribution.

But the more interesting finding lies in the bottom chart: on average the drawdowns under the new scheme are not worse, *but they are not noticeably better either*. Just as with our volatility forecasts, our spot market forecasts do not predict upcoming shocks. Alternative delta hedging acts as an extra "income" generator; it improves returns but should not be viewed as a drawdown control mechanism for VRP strategies.[80]

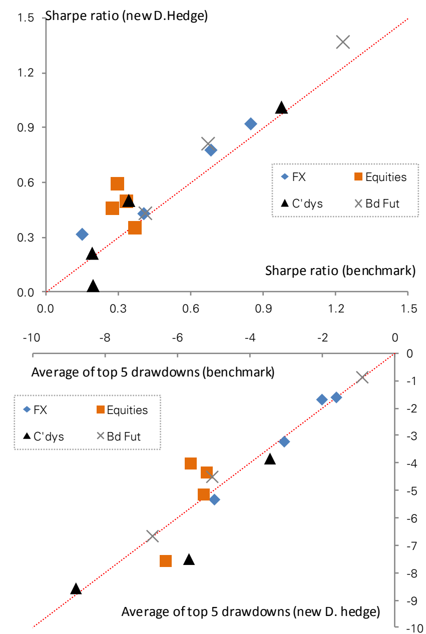
We next evaluate whether these extra returns pollute our capturing of the volatility risk premia. We need to ensure that the marginal gains are not due to structural exposures to static or dynamic factors that a pure VRP strategy does not have. The long run correlations of daily returns between delta hedging P&Ls[81] are at 0.94[82], which suggests we are still capturing the VRP. That said, we need to understand what drives the difference.

In order to search for hidden exposures in each of our 15 asset-specific VRP strategies, we performed an analysis of the attribution of returns. First, we orthogonalised the new delta hedging P&L against the benchmark. We then regressed the residuals against each of 3 explanatory variables: (1) the asset returns, (2) returns from the positions on the asset in our delta-one portfolios from Section 4.1, and (3) the returns from a (time series) momentum strategy on the asset. In our view, these are most likely to incorporate any potential time-homogenous exposures.

Figure 34 shows the time-varying aggregate explanatory power that these 3 variables have on the delta hedge P&L residuals over time, while Figure 35 shows the time-varying betas. The regressions use daily returns, and a 1-year lookback window so as to be adaptive.

We highlight a few observations:

Figure 33: VRP strategy performance – benchmark versus alternative delta hedging



Up to 15 years of returns depending on the market. Source: Deutsche Bank

[80] The ratio of annualised returns to top drawdowns improved in 13 out of the 15 markets under alternative delta hedging versus the benchmark. However, that is mostly due to an improvement in returns and not a reduction in the drawdowns.

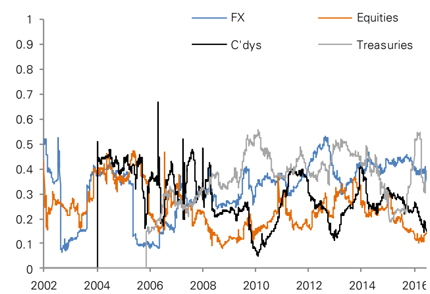
[81] We isolated the delta hedging P&L as opposed to the full strategy P&L in our calculations so as not to bias our correlations upwards. The options mark-to-market is the same for both alternative and benchmark strategies, it is only the delta hedging leg that changes.

[82] We use 15 years of data in FX and most equity markets, and 10-12 years in commodities and Treasuries.

- The regressors explain, on average, 30% of the variations in the residuals over time. *The residuals are not just noise*.
- That said, there is no specific, time-homogenous factor exposure at the aggregate level or specific to a given asset. *All exposures are cyclical*.
- Some asset class-specific patterns have also emerged. In Treasuries, the alternative delta hedge strategy had a long bias, which reflects the multi-decade rally in the asset class. Equities and FX historically loaded positively to asset momentum, except during highly turbulent periods and in the past year – both being instances when markets were notably mean-reverting.

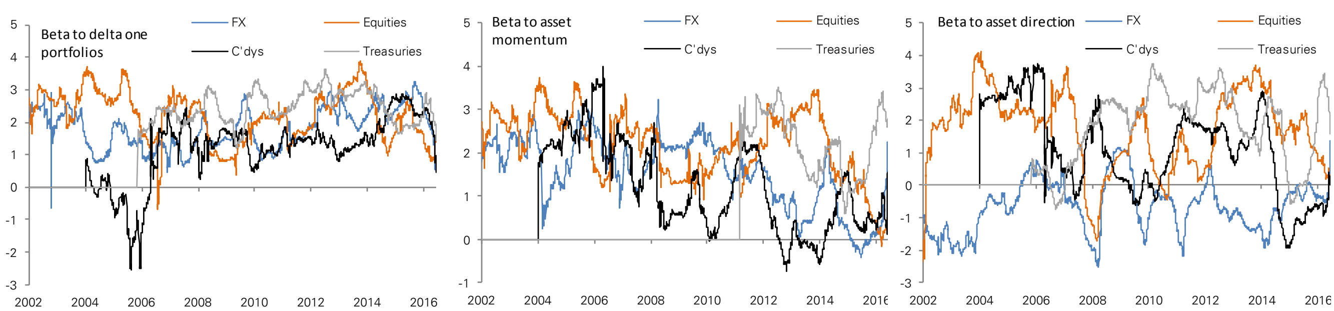
The cyclicality and adaptivity of these results further support the use of alternative delta hedging as a replacement to daily delta hedging. As such, we favour using it when harnessing the volatility risk premium.

Figure 34: Time-varying R-squared – P&L residual regressed against the 3 explanatory variables



1Y rolling window, daily data. Source: Deutsche Bank

Figure 35: Time-varying beta exposures – P&L residual regressed against asset returns, delta one portfolios and asset momentum



Source: Deutsche Bank

6.3 Strategy timing

Timing is the final aspect we will consider in this report.[83] Section 3.2 suggested that exogenous variables may be useful for conditioning exposure during turbulent markets – periods when neither our volatility forecasts nor alternative delta hedging helped. We now delve deeper into this idea.

6.3.1 Timing using the Global Sentiment Indicator

Timing a strategy implies increasing or decreasing the capital allocated to it according to an algorithm. It can be based on endogenous characteristics of the strategy – such as factor momentum or reversal – or defined by variables that are exogenous. We considered timing through 3 indicators: our volatility forecasts, implied volatility and the Global Sentiment Indicator. The first 2 were transformed into 1-year percentile ranks for better comparison versus the third. In each of the cases, we defined the capital allocation ratio as $L_{t,t+h} = 2 - I_t \times 0.02$, where I_t is the level of one of the 3 indicators above and $L \in [0,2]$. The higher the indicator, the more risk averse markets were likely to be, and therefore the lower the capital allocation.

[83] Cross-market diversification and modifying the VRP options basket will be addressed in future reports.

Figures 36 and 37 illustrate the results of timing the original benchmark VRP strategy in each market according to the indicators above. While they correlated heavily to one another, the GSI ranked top in most instances.[84] Using the GSI timer, the drawdowns were cut by approximately 20% from the original and the ratio of annualised returns to average drawdowns almost doubled.[85]

But while these results are encouraging, they are not granular enough. In order to assess how consistent is the improvement across drawdown types, we looked at the ratio of returns between the GSI-timed VRP strategy and the original benchmark in all drawdowns in the latter across all 15 markets. Such analysis is important so as to reduce the sample bias highlighted in Section 3.4.

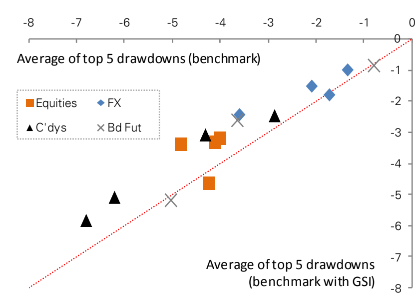
Figure 38 plots the distribution of ratios across 2,185 drawdowns in the aggregated benchmark VRP strategies, bucketed according to the magnitude of loss. Numbers below 1 imply the GSI timer helped, whereas those above 1 imply the opposite. The findings are sobering: on average, *timing only tames medium to large drawdowns* – those above the 30th percentile of the aggregate drawdown distribution. In other words, our proxy for global sentiment is unlikely to capture smaller, idiosyncratic-driven shocks in each particular market, but it should help against stronger, macro-driven drawdowns in the VRP strategy. That said, the wide variations inside each bucket also suggest that the GSI provides no guarantee of improvement.

Figure 36: Sharpe ratio rankings – VRP strategies timed according to 3 indicators

	Implied volatility	Volatility forecast	GSI
S&P 500	3	2	1
SX5E0	3	2	1
Nikkei	3	2	1
Bovespa	3	2	1
E/$	3	2	1
$/Y	2	1	3
A/$	2	3	1
$/BRL	3	2	1
Gold	3	2	1
WTI	2	3	1
Corn	1	3	2
Copper	3	2	1
UST	3	2	1
Bund	1	3	2
JGB	3	2	1
Rank of Average	3	2	1

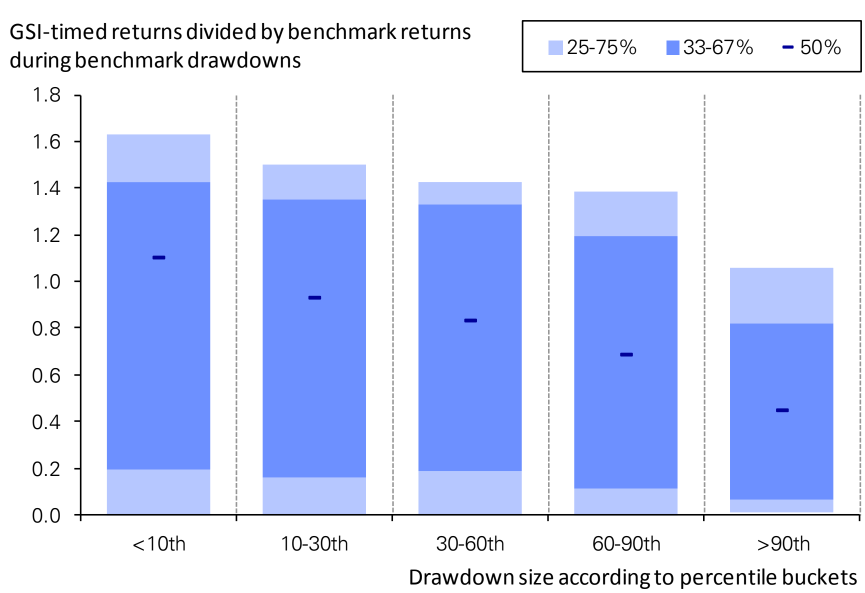
Source: Deutsche Bank

Figure 37: Drawdown profile with and without timing – VRP returns



10-15 years of data depending on the market. Costs are included.
Source: Deutsche Bank

Figure 38: Distribution of drawdown ratios (timed vs un-timed strategy) according to drawdown size



Source: Deutsche Bank

[84] Our volatility forecasts also outperformed the implied volatility in most instances, further confirming the results from Section 4.5.2.
[85] Average of the top 5 drawdowns in each VRP strategy.

6.3.2. Timing using long-term volatility reversal

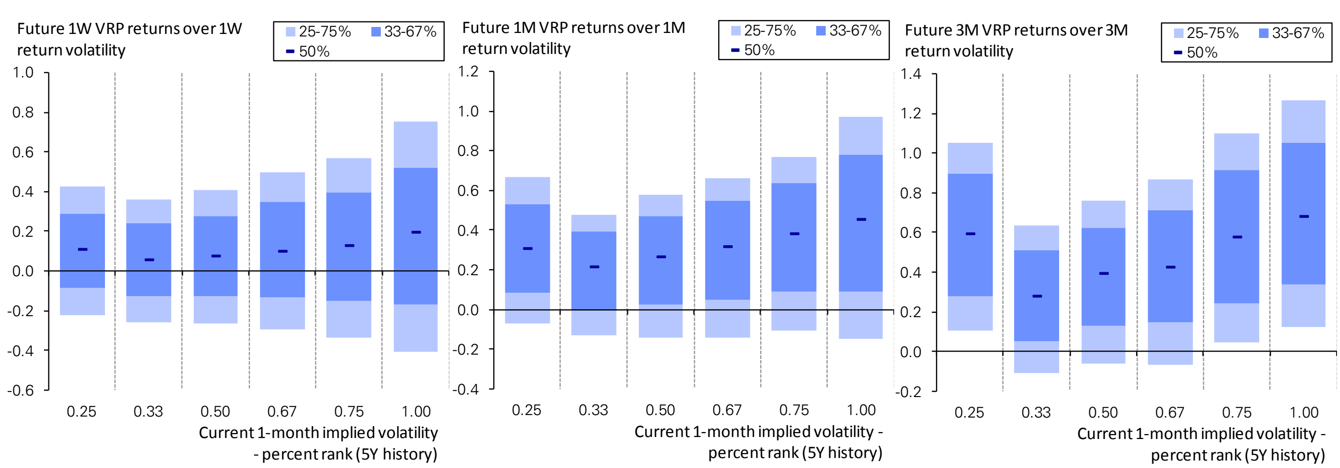
Another, equally pertinent question has to do with the direction of our timing variable. We assume above that rising volatility, and rising risk aversion, should be negative for future VRP returns. However, both implied and empirical volatilities are notably mean-reverting over the long run,[86] which suggests that, at extremes, our actions should be opposite to what is implied by our timing algorithm. In other words, it implies we should lever up the VRP strategy at times of extreme risk aversion and de-lever when volatilities are at an extreme low.

We partly addressed this hypothesis by looking at *future* VRP returns according to *current* implied volatility levels. In order to compare the results across 15 markets, and hence across 15 VRP strategies, we divided future returns by volatility and measured current 1-month implied volatility according to its percentile rank from a 5-year history.[87] We then aggregated all cross-market standardized returns and volatilities into one set.

Figure 39 plots the distribution of future returns according to horizon (1 week, 1 month and 3 months) bucketed according to percentiles of current implied volatility. The results only partly agree with the long run reversal hypothesis described above. Average returns are V-shaped; they do not rise monotonically as implied volatility moves into higher buckets. Further, the pattern becomes more evident as we lengthen the horizon of future VRP returns.

That subsequent volatility-adjusted VRP returns are high at the top volatility quartile is easier to understand, as the strategy is selling volatility when it is at historical highs. But the high (future) returns when (current) volatilities are at the bottom quartile are less intuitive. In our view, this is likely due to anecdotal evidence that volatility stays at low levels for longer than it does at high levels.

Figure 39: Future VRP return distribution according to buckets of current implied volatility



Source: Deutsche Bank

[86] See for, instance, Francq and Zakoian [2010] for a recent reference.

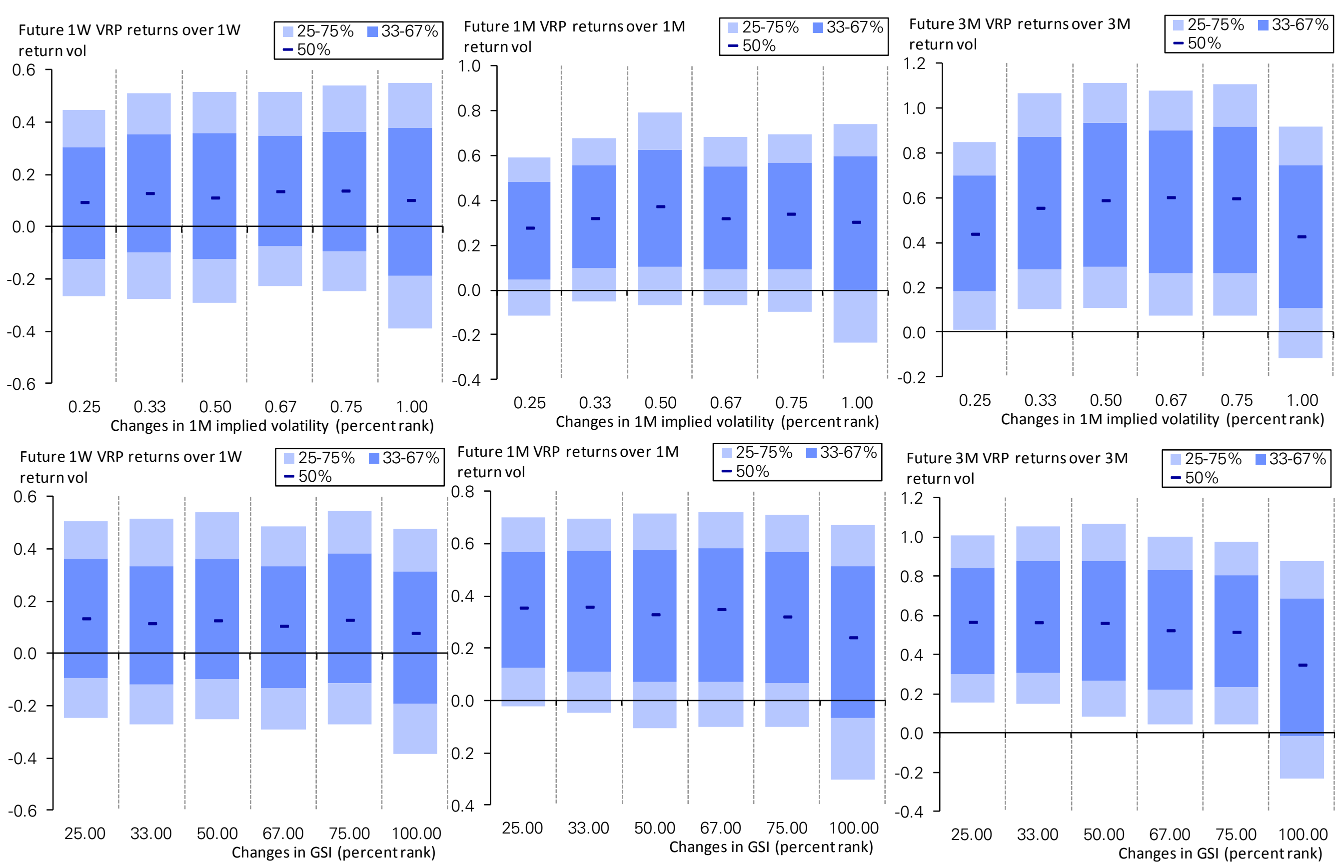
[87] In other words, we take today's implied volatility and calculate where it resides relative to the past 5 years. We do not use future data to estimate the current percentile rank.

6.3.3 Timing using volatility changes

Another important implication from our analysis of long run reversal dynamics has to do with whether this should be incorporated into how we devise the VRP strategy. We have, without testing, opted against that. Increasing capital exposure when volatilities are historically low goes against intuition, and doing the same when they are historically high seems imprudent from a risk perspective. *We rather try to reduce strategy drawdowns than try to speed up its recovery*, as the latter leaves us exposed to potential structural breaks. Further, this V-shaped pattern is not as easily visible when using shorter-dated lookback windows to estimate the volatility percentiles, or when using buckets of our GSI values.

We also looked for patterns in how VRP returns relate to *changes* in implied volatility and the GSI. As before, we standardized and aggregated the changes in *future* VRP returns across markets, and compared these to *current* standardized and aggregated changes in the level of implied volatility and the GSI.

Figure 40: Future VRP return distribution according to buckets of recent change in implied volatility and GSI



Source: Deutsche Bank

Figure 40 illustrates our results, as it plots the distribution of standardized future returns bucketed by 1-month changes in volatility and 1-month changes in the GSI. All charts suggest that *future* strategy returns tend to fall in the upper quartile – when either the current implied volatility or the GSI rose the most versus the recent past. The results are intuitive; significant jumps in volatility and the GSI imply risk aversion, which is detrimental to future VRP returns.[88] These findings indicate the need for further research, as they alone are not enough to justify changing the format of our timing algorithm.

6.4 Timing or Alternative Delta Hedging?

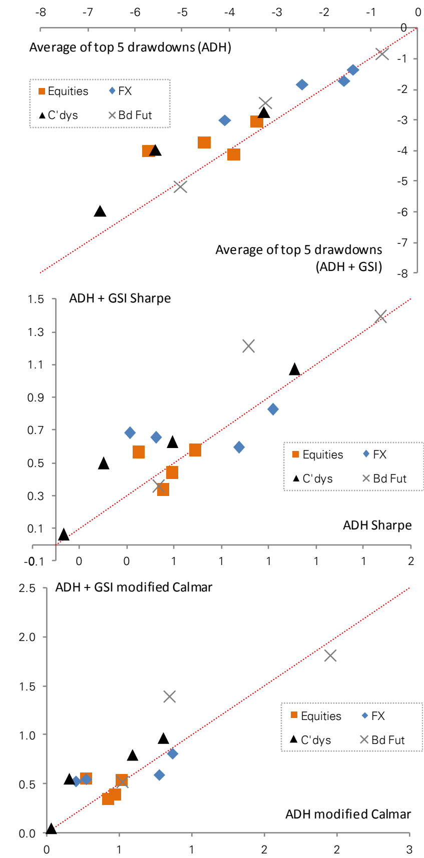
So far we have shown two VRP-enhancing approaches: alternative delta hedging (ADH), introduced in Sections 5.4 and 6.2, and timing, introduced in Section 6.3. The former acts as an "income" provider, whereas the latter reduces drawdowns. A final question is whether we should combine both.

Figure 41 compares the returns of the ADH strategy on its own versus the ADH strategy combined with timing. The drawdowns improved in most cases - as expected - but the resulting Sharpe ratios and modified Calmar ratios[89] are better in only half of the markets. Timing mechanisms often curtail income, as they can lead to under-leveraging during recovery periods, and timing the VRP seems no different.

Figure 42 delves deeper into this topic. We zoom into the exact difference between the ADH strategy and the benchmark - the P&L difference between both delta hedging legs - and evaluate whether there is value in timing that spread.[90] As the Sharpe ratios show, there is no strong evidence that timing the spread adds value, just as we find no evidence that our aggregate delta one portfolios from Section 4.1 can be timed. The results look particularly worse in Equities, as the orange boxes in Figure 42 show.

In summary, these results show there exists a trade-off between income boosting and drawdown control when trying new VRP enhancement methods. While we favour alternative delta hedging in all instances, we believe the decision to further apply strategy timing should be dependent on the investor's risk constraints.

Figure 41: Performance comparison – ADH alone versus ADH + timing



Source: Deutsche Bank

6.5 An Equity VRP Strategy

We finalise this section by showing our results in 3 equity indices: S&P 500, Eurostoxx 50 and Nikkei 225. In all 3 cases, we systematically sell USD 100 worth of a basket of 1-month options, rolled every month. The basket is comprised of 10-delta, 25-delta and ATMF calls and puts, where the capital allocated is inversely related to the strike level - as per variance swap formula. As elsewhere in Section 6, costs are included.

[88] The reader will likely see another, smaller pattern in the form of an inverted u-shape in the chart pertaining to future 3-month returns against 1-month changes in implied volatility. It suggests that significant drops in volatility are also detrimental to future VRP returns. It suggests, in fact, that the most favourable environment is one where implied volatilities are not moving (the 50-67th percentile bucket).

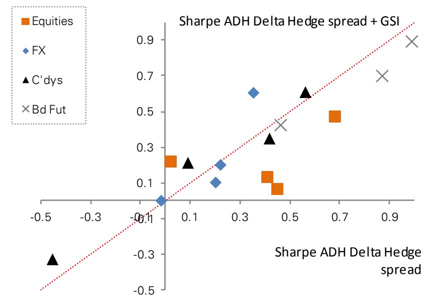
[89] Annualised strategy returns divided by the average of its 5 worst drawdowns.

[90] We effectively divided the strategy P&L into 3 legs: the options P&L, the benchmark delta hedge P&L and the additional delta hedge P&L from the ADH approach. We focus only on the third leg because we have already shown earlier that the GSI performs well in timing the other 2.

Figure 43 and Figure 44 show our backtested results. The benchmark strategy applies daily delta hedging, and is plotted in gray. The strategy that applies the GSI indicator to adjust position sizes in the benchmark strategy, as described in Section 6.3, is plotted in blue. The strategy that delta hedges according to our expected spot returns, as described in Section 5.4, is shown in blue. Finally, the strategy that uses both alternative delta hedging and timing is plotted in orange.

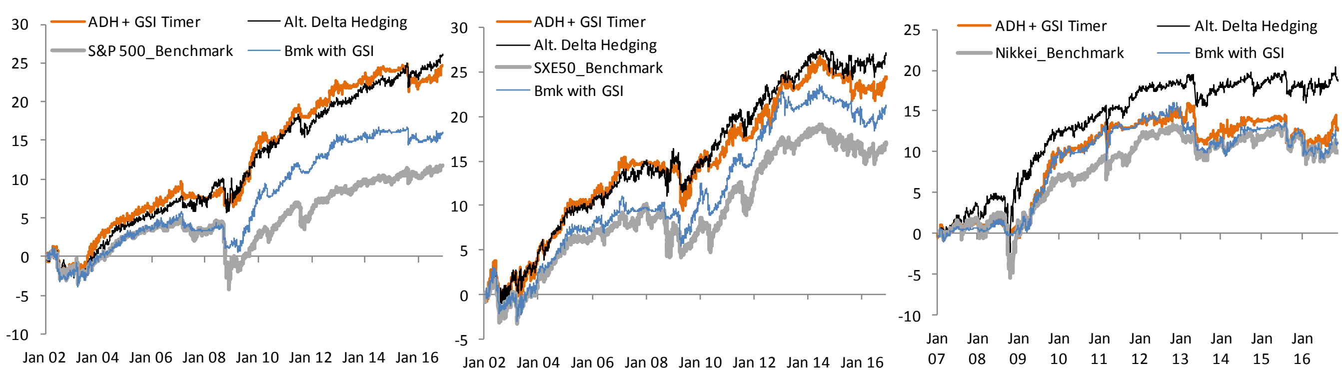
Alternative delta hedging boosts the returns in all 3 cases; it bumps the slope of our 3 time series. The GSI timer reduces the average of the 5 worst drawdowns in 2 out of 3 strategies. When combined, the results are also mixed – on average, we see lower risk-adjusted returns but lower drawdowns.

Figure 42: ADH delta hedge spread with and without timing



Source: Deutsche Bank

Figure 43: Backtested cumulative returns – VRP strategy applied to equity index options



Source: Deutsche Bank

Figure 44: Backtest performance characteristics

	Benchmark			Alternative Delta Hedging			GSI timer			ADH + GSI timer		
	S&P 500	SX5E	Nikkei	S&P 500	SX5E	Nikkei	S&P 500	SX5E	Nikkei	S&P 500	SX5E	Nikkei
Ann. Ret	0.8	1.1	1.0	1.7	1.8	1.9	1.05	1.38	1.11	1.6	1.6	1.3
Ann. Vol	2.7	3.4	3.7	3.0	3.7	4.2	2.51	3.31	3.34	2.9	3.7	3.8
Sharpe Ratio	0.30	0.33	0.28	0.59	0.49	0.45	0.42	0.42	0.33	0.57	0.44	0.33
Sortino Ratio	0.34	0.42	0.32	0.68	0.64	0.53	0.50	0.54	0.39	0.67	0.55	0.37
Avg Top 5 DDs	-4.0	-4.2	-4.1	-3.4	-3.9	-4.5	-3.20	-4.65	-3.34	-3.0	-4.1	-3.7
Length Top 5 DDs (yrs)	1.6	1.7	1.2	0.9	1.1	0.8	1.7	1.9	1.1	1.3	1.5	1.1
Mod. Calmar	0.2	0.3	0.2	0.5	0.5	0.4	0.3	0.3	0.3	0.5	0.4	0.3

Source: Deutsche Bank

7. Conclusion

This report introduces a framework for extracting value in volatility surfaces across asset classes. It is based on comparing the market-implied distribution of future returns - the Q-distribution - with our subjective expectations of the same - in other words, our P-distribution. In building the latter, we use a parametric approach to combine the information from our delta-one systematic strategies with a regime-switching, event-calibrated multivariate risk factor model.

Today's *Derivatives Spotlight* focuses on the volatility parameter of the distributions above. More specifically, on the volatility risk premium, a primary source of value in options trading. We describe the merits and shortcomings of popular VRP enhancement methods, and introduce a different approach based on volatility forecasting and alternative delta hedging.

We compare our methods to multiple benchmarks focusing on both statistical accuracy and strategy returns. As is often the case with rigorous testing, the results were not always as encouraging as we had hoped. That said, we believe that the final recommendations - based on alternative delta hedging and timing - should improve the way investors harness this rich source of risk-adjusted returns.

Our conclusions are as follows:

- Having some predictive power of future asset returns – such as when using our real world P-distribution – can be valuable in improving VRP strategy returns.
- In our context, the first moment of this distribution has been key in improving the way we *delta hedge*. That improvement came from higher risk-adjusted returns, although not from lower drawdowns.
- For the latter, *timing* has shown some encouraging results. We proposed a method based on our Global Sentiment Indicator – which adapts fast to changing market conditions.
- Future research will focus on other uses of the P-distribution and how we can use it to identify and extract more sources of value in the volatility surface.

Appendix I: Our Multi-Factor Risk Model

This section introduces our multi-factor risk model for cross-asset returns. While its output can be used for multiple purposes in the future, here we use it as input to the asset volatility forecasts introduced in Section 4. Our model is based on the Arbitrage Pricing Theory (APT) – an extension of the Capital Asset Pricing Model (CAPM) which evaluates the sensitivity of asset returns to a set of risk factors. [91]

The risk factors we use encompass macro drivers (inflation and growth), market drivers (sector and asset class returns) and dynamic drivers (Momentum and Carry). We calibrate the model to 80 assets across currencies, equity indices, commodities and international Treasuries.[92] Asset sensitivity to our global drivers is estimated through *stepwise robust regressions*.

Section I.I provides a background to risk factor models. Section I.II provides the algorithmic framework for setting up the factors and volatility forecasts. Finally, Section I.III focuses on model estimation – in other words, how our betas are derived.

I.I Background and Choice of Framework

Multi-factor models of asset returns, a.k.a. multi-factor risk models, are commonplace in quantitative investing. They serve 3 primary purposes: dimensionality reduction, better attribution of returns and better estimation of how factors and assets interact. The first is key to large institutional portfolios; instead of evaluating the covariances between thousands of securities, one can model less than 100 factors instead – a much more palatable exercise. Risk attribution is equally important; factor models decompose portfolio risk into a smaller, manageable list of sources, making it easier to manage and protect positions. Finally, risk models account for the interaction between factors and assets, thus addressing spill-overs and other effects that are missed out by univariate models. Combined together, these 3 features provide risk models with an edge when predicting future asset volatility – our ultimate goal.[93]

Risk models are normally divided into 3 types: macroeconomic, fundamental and statistical:[94]

- Macro factor models use observable economic anchors – tradable and non-tradable – as risk drivers. The data pre-specifies the return of each

[91] See Roll and Ross [1976] and Burmeister et al [1994].

[92] 21 equity index futures: ASX, Bolsa, Bovespa, CAC, DAX, Eurostoxx, HSI, IBX, ISE, JSE, Kospi, Nasdaq, Nikkei, OMX, RDX, SMI, S&P 500, TSE, TWE, FTSE and WIG. 30 USD/FX: USD/G10 + USD vs BRL, CLP, COP, CZK, HUF, IDR, ILS, INR, KRW, MXN, MYR, PEN, PHP, PLN, RON, RUB, SGD, THB, TRY, TWD, ZAR. 20 commodity futures: brent, cocoa, coffee, corn, cotton, gasoil, heating oil, natural gas, soybeans, sugar, wheat, WTI, silver, aluminium, gold, copper, led, nickel, platinum, zinc. 9 bond futures: 10Y government bonds in Australia, Canada, Switzerland, Germany, UK, Japan, Mexico, New Zealand, US.

[93] See Ward [2010], Ward et al [2016] and Grinold & Kahn [1999] for an in-depth overview of risk factor models and risk modelling.

[94] See Zivot [2011] and Connor [1996] for more details on these 3 types.

factor, and its loading to an asset is estimated through time series regression techniques. Macroeconomic factor models are the simplest and most intuitive of the 3, though suffer from 2 potential drawbacks: *mis-specification* (when we use the wrong variable to represent the factor) and *factor omission* (when we ignore a key driver).

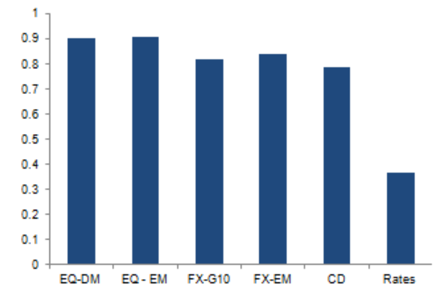
- Fundamental factor models are distinct in that we assume that the factor loadings are pre-specified, and factor returns are estimated through cross-sectional regressions. They are the standard followed by risk factor providers, and use asset-specific characteristics such as industry sector, corporate accounting metrics and other style classification measures. While popular in the equity investment community, this type of factor model is less applicable to markets with few observable asset-specific characteristics. It also suffers from the same drawbacks as macro factor models, though have the advantage of being memory-less – it allows for *point-in-time* risk estimation.
- Statistical factors take a completely different approach; neither the factor returns nor their loadings are observable. Both are estimated instead through statistical techniques – most commonly, factor or principal component analysis. Mis-specification and factor omission are not a concern in this approach, but statistical factors suffer from a lack of direct economic interpretation.

Choosing from the 3 alternatives is not straight-forward; the researcher must take a view on whether she needs the factors to be interpretable, whether there are enough asset-specific characteristics that serve as drivers, whether there are enough assets for cross-sectional regressions to be run efficiently, and – ultimately – what she needs the model for.

In our case, we use it to forecast the volatility of asset returns. With that in mind, the macro factor model approach suits us best. Our pool of assets is too small for cross-sectional regressions to be reliable, and our assets lack enough fundamental characteristics. At the same time, we want to be able to interpret what each risk factor represents for future purposes.[95]

Figure 45 illustrates part of our argument – that the fundamental factor model approach does not fit our goals as well. It plots the correlation between factor returns *observed* by market data – the “true” returns – and factor returns *estimated* by the cross-sectional regressions defined using the fundamental factor approach.[96] The correlations are generally strong but not always; in the case of rates markets, our estimated returns correlate little to the true factor returns due to a lack of breadth of constituents.

Figure 45: Correlation between observed and estimated factor returns



Source: Deutsche Bank

I.II Selecting our Risk Factors

As the reader may suspect, selecting the qualifying factors was the biggest challenge we faced when creating our risk model. We first had to define how much of the variance in our basket of assets needed to be explained by

[95] Examples of future use include portfolio risk hedging and risk factor investing.

[96] We classified our pool of 80 assets into 6 sub-asset class buckets: DM and EM equities, DM and EM FX, commodities and international Treasury markets. We assigned a value of 1 or 0 to an asset depending on whether it belonged to that bucket, in the same manner as stocks are classified into industry or sector buckets. We also added a global market bucket which applied to all assets. We then applied constrained OLS regressions to estimate factor returns.

common factors, and then we had to choose which factors best fit the task. Both steps are challenging. The target explanatory power is a random variable in itself, as the interaction between factor and idiosyncratic risk changes over time. Choosing the actual factors is not straight forward either. We must acknowledge our hidden bias – we know which factors explained the past – and seek to minimize its adverse effect on future forecasts.

I.II.I Defining the explanatory power target

The first task is to define how much of the variance in our asset pool should be covered by the multi-factor model. In other words, we must define a threshold between factor risk and idiosyncratic risk. To do so, we resorted to Principal Component Analysis (PCA). This method can differentiate between factor and idiosyncratic variance as it outlines the marginal explanatory power of each factor and how it loads into each asset.

There are 2 aspects to this task: defining the maximum number of statistical factors, which is defined once using a long historical window, and choosing the optimal number of factors applicable to a given rebalancing date. We applied the *Cattell's Scree* test to define the former, choosing a maximum number of 8 (out of 80) statistical factors.[97] As for the latter, we applied the information criteria proposed by Bai and Ng [2002].

Next, we want to evaluate how much of the variance in each asset can be explained by the statistical factors chosen above. We run standard time series regressions of asset returns on principal component returns and aggregate the R-squared output from the 80 individual regressions.[98] This number, which changes at every rebalancing date, becomes our target explanatory power. It is the amount of variance we will try to explain using the macro factors built below.

I.II.II Choosing our factors

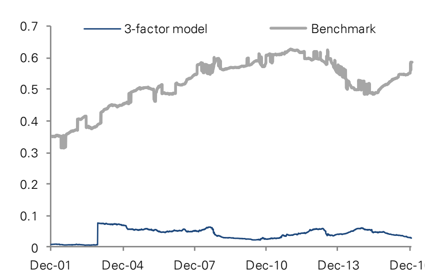
The premise of a macroeconomic risk factor model is, unsurprisingly, that asset returns are chiefly driven by macroeconomic developments. As such, it is a sensible model choice for our markets given that they reflect the typical investment scope of a macro investor.

This choice, however, creates 2 challenges: how to capture the desired factors, and whether we are capturing enough of them.

We started with Chen et al [1986]; as per authors, we assumed that our asset returns should be explained by surprises in GDP, inflation and interest rates. We tried capturing the first 2 through the *nowcasting* indices introduced in Natividade et al [2015], and the latter through 10Y US Treasury returns.

Figure 46 compares the average explanatory power achieved through these variables versus the target defined in Section I.II.I, making it clear that we have not captured enough. It points us back to the 2 challenges highlighted above;

Figure 46: Explanatory power using *nowcasting*-based macro factors and interest rates versus the target



Source: Deutsche Bank

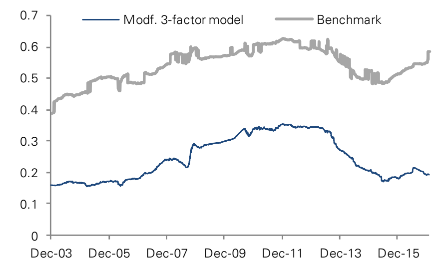
[97] The *Cattell's Scree* test plots the components as the X-axis and the corresponding eigenvalues as the Y-axis. The contribution of each eigenvalue in explaining the total variance is presented in descending order and linked with a line. Once the drop ceases and the additional eigenvalues explain little extra, we define the cut-off. It suffers from the drawback that there is no deterministic solution as to the optimal number of factors; the choice is subjective to the researcher. See Cattell [1978].

[98] The OLS regressions were run on 2 years of daily data, rolled daily.

nowcasting indices may be of higher frequency but perhaps still too slow, and we may be missing other relevant drivers.

Next we tackled the factor representation issue. Market returns are often cited as representative of macroeconomic developments[99], and as such we replaced our growth indices with asset class returns[100] in G10 equities, G10 USD/FX, commodities and global Treasuries. These should capture both the macro picture and asset class specific innovations, and correspond to the "market" variable alluded to in the CAPM. We also replaced our inflation series with a market-derived metric: long nominal 10Y USTs while short 10Y TIPS, as per Podkaminer [2013]. Figure 47 shows the effect of the new set on explanatory power; we are capturing more, but still not enough.

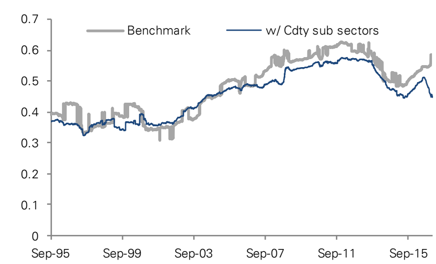
Figure 47: Explanatory power using market-based macro factors versus the target



Source: Deutsche Bank

Our next step was to either add market-representative categories or to sub-categorise our factors in search for additional entropy. This is akin to industry and sector classification in fundamental equity risk models. In the case of equities and FX, we added emerging market return factors by orthogonalising EM asset returns against G10 asset returns. In commodities, we split the market factor into 3: energy, metals [101] and agricultural asset returns. Separating between developed and emerging markets is not uncommon – the latter often captures stronger country risk premium – see, for instance, Podkaminer [2013] and Greenberg et al [2016]. Figure 48 shows the pickup in explanatory power.

Figure 48: Explanatory power using sub-categorised, market-based factors versus the target



Source: Deutsche Bank

Finally, for completeness, we added two dynamic factors: time series Momentum and Carry. While these base factors are typically classified as *return* drivers instead of *risk* drivers, there will be periods when they drive portfolio losses in ways not captured by the market factors outlined earlier. The attractive long-term returns in Carry and time series Momentum investing exist so as to compensate for these specific short-term losses.[102] The procedure is consistent with Ang [2014].

Figure 50 outlines our final list of risk factors.[103] We acknowledge a bias; we assume these factors, which "worked" in the past, will also "work" in the future, but our consistency with the literature is reassuring. Figure 49 shows the explanatory power of our final multi-factor risk model; it looks very similar to the original, PCA-driven target set in Section I.II.I. in other words, we can now explain the desired amount of variance in our pool of 80 markets using "tangible" – and even hedge-able – drivers.

[99] See Ang [2014] for a detailed discussion.

[100] We proxied asset class returns through (first) principal component baskets. The PCs are estimated using a 1-year lookback window of daily returns, rebalanced monthly. The factors are extracted from a correlation matrix – i.e. we assume unit variances.

[101] We chose not to separate between base and precious metals because of the limited breadth in the latter. Had we sub-categorised, the regressions run on their 1st principal components would have over-estimated the explanatory power of the model.

[102] The "momentum crash" effect is a good example.

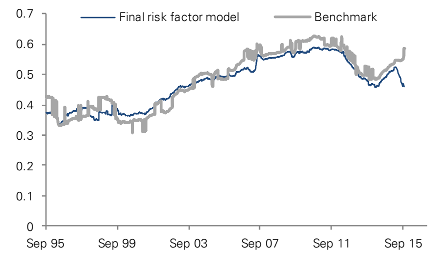
[103] The reader may also be interested in knowing that the only additional factor considered was "crash risk", introduced in David and Bhansali [2014] and proxied as the returns of a short 10-delta 1-month put on the S&P 500. We ultimately opted against it because it failed to explain much extra variance.

Figure 50: Our factors - description and how to attain exposure

Risk factors	How to get exposure
Inflation	Return of long nominal US Treasuries, short TIPS (Treasury Inflation-Protected securities) portfolio
Equity – Developed Market	Broad-developed market equity index return. This is gained through the returns of 1st principal component of our developed market indices.
Equity – Emerging Market	Return of long EM equity index (1st principal component of emerging market indices), short DM equity index portfolio.
Dollar – G10 countries	Long 1st principal component of USD/G10 currencies.
Dollar – Emerging Markets	Return of long 1st principal component of USD/EM currencies, short 1st principal component of USD/G10 currencies portfolio
Commodities – Energy sector	Long 1st principal component of our energy basket.
Commodities – Metals sector	Long 1st principal component of our global metals (industrial and precious) basket.
Commodities – Agriculture sector	Long 1st principal component of our agriculture basket.
Rates – Global Treasuries	Long 1st principal component of our sovereign 10Y bond futures market.
Cross-asset Trend	Return of our cross-asset trend strategy (Marta).
Cross-asset Carry	Return of our cross-asset carry strategy (Carrie)

Source: Deutsche Bank

Figure 49: Explanatory power of our final risk factor model versus the target



Source: Deutsche Bank

I.III Model Estimation

Having defined our ingredients in Section I.II, we now move into the final step: how to estimate the model. Our final goal is to estimate the sensitivity of each asset to our list of risk factors, as these factor loadings will serve as input to how we forecast asset volatility in Section 4.

As is commonplace in quantitative research, we first need to address two natural issues with market data: the interdependence between our factors, and data outliers. More specifically:

- We addressed *multicollinearity* risk in 3 ways. First, we *orthogonalised* some of the variables; that is, we regressed EM equity and FX returns on G10 equity and FX returns, and used the respective residuals as our emerging market factors. Second, we ran the Belsley Collinearity Diagnosis[104] on a rolling basis to evaluate the degree of collinearity between our explanatory variables. The results were largely satisfactory.[105] Finally, as a means of protecting ourselves against the threat of collinearity in the future, we applied *stepwise regressions* when estimating asset sensitivity to each factor. Through this procedure, only relevant factors were used at each estimation date; insignificant factors were dropped.[106]
- We sought to neutralise outlier influence in 2 ways.[107] First, we removed the edges of our factors returns (1st and 99th percentiles). Second, our

[104] See Belsley and Kuh [1980] for details.

[105] We used 17 years of data in total. The tests were run using 2 years of daily factor returns, rolled daily. The test statistic did not breach the critical threshold value of 10 at any point of our sample history.

[106] We also attempted more complex techniques, namely: partial least squares regressions, principal component regressions, ridge regressions and LASSO. None showed enough improvement in our results so as to justify the extra number of parameters required for calibration.

[107] Outliers often lead to a violation of the assumption that our regression residuals are Gaussian, prompting issues such as heteroskedastic and skewed residuals.

stepwise regressions were conducted in *robust* form – in other words, through robust stepwise regressions.[108]

At every rebalancing date, the steps of our algorithm are as follows:

1. We first take the last 2 years of daily observed returns in both asset and factor data. We de-trend and remove the outliers[109] in both cases.

2. We regress asset returns against each factor using a robust regression approach:

   i. We first define w as a *N*x1 unit vector, where *N* is the number of data points, and $W = \left(uw^T\right) \bullet \mathrm{I}$ where u is a *N*x1 unit vector, I is the *N*x*N* identity matrix and $\bullet$ is the dot product. W is therefore a diagonal matrix whose diagonal elements equal the elements of w.

   ii. We estimate the sensitivity of asset returns to each factor using a weighted least squares regression, i.e. $\beta = \left(X^T W X\right)^{-1} X^T W Y$, where *Y* is a *N*x1 vector of asset returns, *X* is a *N*x13 matrix of factor returns (plus the intercept), and β is a *13*x1 vector of coefficients.

   iii. We redefine $w = \left(|s| < 1\right) \bullet \left(1 - s \bullet s\right) \bullet \left(1 - s \bullet s\right)$. Further:

   $s = \frac{r}{t\sigma\sqrt{1-h}}$, where r is a *N*x1 vector of residuals from the regression in ii, $\sigma = {MAD}/{c}$, $t = 4.685$ and $c = 0.6745$ are constants[110] MAD stands for mean absolute deviation and $h = diag\left(X\left(X^T X\right)^{-1} X^T\right)$ representing a vector of leverage values from the current regression fit.[111]

   iv. Repeat steps i – iii until w is stabilized. It represents our *robust weights.*

3. We run *stepwise* weighted regressions of asset returns on each of the risk factors using the robust weights estimated in Step 2:

   i. We run univariate (weighted) regressions on each risk factor and choose the factor whose P-value is the smallest, assuming it is also below a qualifying threshold.

   ii. We re-run our regressions on every remaining factor, having already included the term that qualified above, and keep the next relevant factor – assuming the P-value qualifies.[112]

---

108 The method is called iteratively weighted least squares. It involves iteratively re-weighting each data observation and re-running our regressions through weighted least squares until our betas converge to a target tolerance level. We use a bi-square weighting function to re-weight our data observations. See DuMouchel and O'Brien [1989] for details.

109 We remove observations below the 1st percentile and above the 99th percentile of daily returns over the past 2 years.

110 Setting *t* = 4.685 gives us coefficients that are 95% as statistically significant as the OLS estimates. Setting *c* = 0.6745 makes the estimate unbiased in a normal distribution.

111 See, for instance, https://en.wikipedia.org/wiki/Leverage_(statistics)

iii. We repeat the step above until there are no qualifiers left. For every asset, we record the qualifying factor coefficients and set the others to zero.

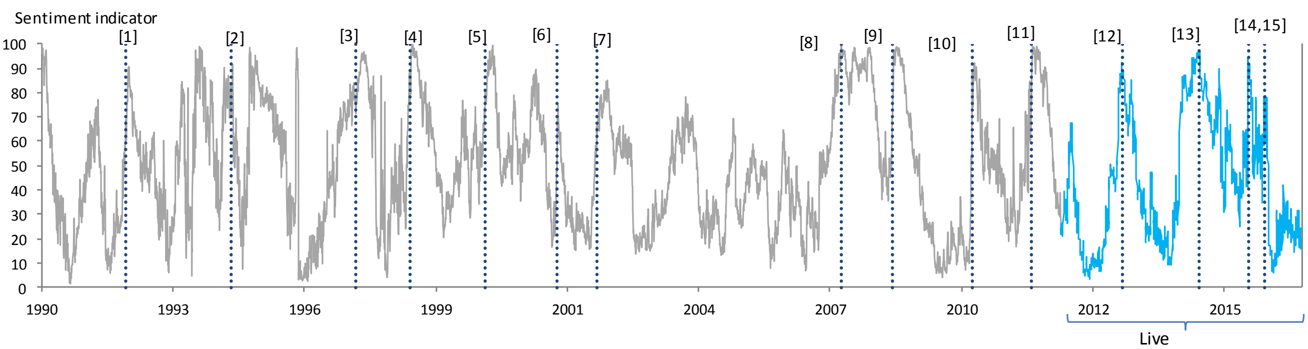
The coefficients defined in Step 3 are used as input to our asset volatility forecasts, as defined earlier in Section 4.3. This completes the estimation of our cross-asset, macro factor risk model.

[112] If the p-value of our regression F-statistic with the new explanatory variable is higher than that of without it, we do not include the new variable. For reference, see: https://uk.mathworks.com/help/stats/stepwiselm.html?searchHighlight=stepwiselm&s_tid=doc_srchtitle

Appendix II: Our Global Sentiment Indicator

Modeling investor risk appetite has always been a popular topic in both academic and industry research. In March 2012, we – Chen and Natividade [2012] – introduced Deutsche Bank's Global Sentiment Indicator (GSI), a variable that points to the current state of market sentiment. It is available on Bloomberg as *DBQSGSI Index*. Figure 51 shows its historical performance prior to and after publication.

Figure 51: Global Sentiment Indicator time series

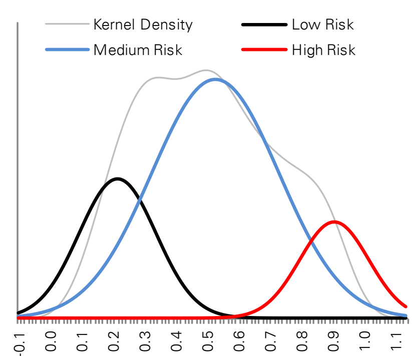


[1] Black Wednesday, [2] Mexican devaluation, [3] Asian crisis, [4] Russian crisis, [5] Dot-com bubble, [6] 9/11, [7] Market crash, [8] Credit crunch starts, [9] Lehman Brothers, [10] Greek downgrade, [11] European debt woes, [12] Taper Tantrum, [13] Greek debt woes, [14] Oil Price Shock, [15] British EU Referendum. Source: Deutsche Bank

The GSI aggregates 11 barometers of risk across asset classes. These were collected through a survey of our colleagues in macroeconomic research and focused on variables that were highly adaptive and historically persistent – in other words, they did not predict turbulent periods but adapted quickly to them. They are:

- Equity implied volatility: the VIX, a weighted average of implied vols which helps gauge the market expectation of how volatile the S&P 500 will be over the next 1 month.
- Financial sector risk (equity perspective): the ratio of the MSCI Financials Local Index over the MSCI World Index.
- Financial sector risk (rates perspective): the difference between 30Y and 2Y asset swap spreads, as the former represents possible stress in the pension fund and insurance sector while the former represents banking sector stress. This measure of liquidity risk has been more efficient at market crises prior to the most recent.
- Interest rate implied volatility: the average of USD, EUR and JPY 3M5Y swaption volatility.
- Short-term interest rate liquidity risk: TED spread, the difference between interest rates on interbank loans and US T-bills.
- Investment grade credit spreads: the spread between Moody's BAA corporate bond index and 10Y US Treasuries.
- Non-financial CDS spreads: the iBoxx US Non-Financials Index.

Figure 52: GSI – Kernel density and Gaussian mixture decomposition



Source: Deutsche Bank

- FX implied volatility: the CVIX, a turnover-weighted average of 3M implied volatility which helps gauge the market expectation of how volatile currency markets will be over the next 1 month.
- FX volatility slope: the average term spread between 1Y and 1M volatility in EUR/USD, USD/JPY and EUR/JPY.
- FX volatility skew: the average level of 1M 25-delta risk reversals between USD/JPY, EUR/USD and EUR/JPY.
- EM sovereign risk: JP Morgan's EMBI+ index.

We standardized the data by calculating the rolling 1-year percentile of each series, and distilled the output through its first principal component.[113] The final variable is bounded between 0 and 100, where higher readings imply a higher risk environment.

A Gaussian mixture decomposition of the GSI's empirical distribution pointed to the existence of 3 distinct regimes of risk appetite – low, medium and high risk – as consistent with its tri-modal kernel density. Over the past 27 years the GSI has spent roughly 33% of the time in low risk, 40% in the intermediate regime and the remaining in high risk. Figure 52 plots this historical distribution.

Figure 53 illustrates how efficient the GSI has been at capturing turbulent periods in both backtest and live history. In 13 out of the 17 stress periods (and 4 out of 4 live stress periods) the indicator has been at its upper tercile, suggesting it adapted to high risk conditions when needed. It did not predict any particular stress period, as it is not designed to do so, but adapted quickly to each of the periods below.

We have used the GSI for 2 purposes: *timing* risk-sensitive strategies and estimating regime-dependent covariance matrices:

- *Timing* a strategy implies increasing or decreasing the capital allocated to it according to an algorithm that is often based on values of an exogenous variable. In this and prior reports, we have used GSI levels to increase or decrease exposure according to the ratio: $L_{t,t+h} = 2 - I_t \times 2$, where $L \in [0,2]$. Figure 54 shows the rolling drawdowns on a standard FX Carry strategy (*DBHVBUSI Index*) before and after we applied the algorithm above, over the long run and since we published the idea in Anand et al. [2014]. Risk indicators, when applied to risk-sensitive strategies, provide rare instances when timing is promising. On average, timing is a very hard task.[114]

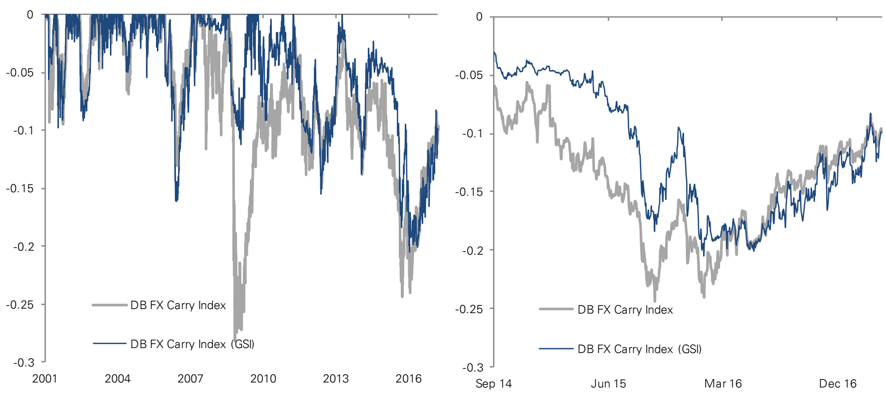
Figure 53: GSI levels across historical risk events

Event	Start	End	Average GSI Level
UK withdraw from EUR. ERM	Jul-92	Oct-92	41
US rates shock/Orange county	Feb-94	Jun-94	77
Asian Financial Crisis	May-97	Feb-98	69
RUS default	Jul-98	Sep-98	59
.COM bubble	Feb-00	Jun-00	79
.COM bubble	Oct-00	Nov-00	65
.COM bubble	Feb-01	May-01	74
2002 downturn	Jul-02	Oct-02	72
Start of US hiking cycle	Mar-04	Jul-04	65
Start of financial crisis	Jun-07	May-08	89
Collapse of Lehman	Sep-08	Jan-09	88
EU Sov. Debt	May-10	Jul-10	82
EU Sov. Debt	Jun-11	Dec-11	80
Taper Tantrum	May-13	Sep-13	72
Greek Debt Crisis	Sep-14	Jul-15	79
Oil Price Stress	Feb-16	Mar-16	79
British EU Referendum	Jun-16	Jul-16	68

Source: Deutsche Bank

[113] We set negative sample covariance values to zero, as per Luo et al [2009]. All data is pre-processed such that it has the same direction to risk appetite, therefore minimising the instances where we get negative covariances in the first place.
[114] See Asness [2016] for a longer discussion of this topic.

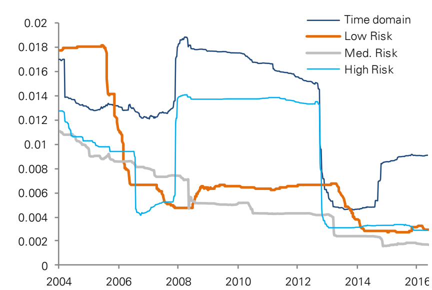
Figure 54: Rolling FX Carry (DB Balanced Harvest) drawdowns – with and without GSI timing, historical and since September 2014



Source: Deutsche Bank

- *Regime-based co-movement estimation* is another area in which the GSI can be applied. Quantitative researchers often calculate co-variances using asset returns sequenced in time. Another approach, however, is to estimate co-movements conditioned on data of the same market regime over different – potentially non-sequential – periods in history. The premise is that asset dependencies may be better modeled as a function of risk states; they may, for instance, be more strongly correlated in risk aversion as one factor – "risk" – becomes the primary driver of asset returns. Conversely, they may be less correlated under lower risk conditions as multiple drivers affect price action. Figure 55 plots the volatility of the average pairwise correlation between 18 cross-asset underlyings, bucketed according to 3 GSI states and according to time.[115] It shows that the regime-conditioned estimates are less volatile than the estimates calculated in time domain; which suggests the former could potentially lead to more stable and adaptive asset weights when building portfolios.[116]

Figure 55: Volatility of the average pairwise correlation using different (rolling) correlation matrices



Source: Deutsche Bank

In summary, the GSI seeks to update the investor about current market conditions in as adaptive way as possible. It can be a powerful tool for modeling financial data – either for timing or estimating co-movements.

[115] We used 4 flagship assets in FX, commodities and global Treasuries, and 5 equity indices. In this exercise, we effectively applied a bucketing approach to estimate joint relationships. First, we estimated the 3 risk-state dependent correlation matrices by separating asset returns into 3 buckets based on the GSI level at the time. In this case, we used an anchored lookback window so as to maximise the number of data points. In the time domain, however, our correlation matrices were estimated using a lookback window equal to the average size of 3 risk-state buckets. As such, the number of data points in each bucket rises in time equally in both domains. Thereafter, we calculated the volatility of average pair-wise correlation over time as a measure of the stability.

[116] Researchers then control for covariance estimation error using shrinkage or factorization. For shrinkage, see James and Stein [1961], Jorion [1986] and Frost and Savarino [1986]. For factorization, see Sharpe [1963], Chan et al [1999] and MacKinlay and Pastor [2000].

8. Bibliography

Ahmad, R. and Wilmott, P. [2005], Which Free Lunch Would You Like Today, Sir?: Delta Hedging, Volatility Arbitrage and Optimal Portfolios, Wilmott Magazine (November 2005): 64-79.

Aït-Sahalia, Y., & Mancini, L. [2008]. Out of sample forecasts of quadratic variation. Journal of Econometrics, 147(1), 17-33.

Alexander, C., & Lazar, E. [2006], Normal mixture GARCH (1, 1): Applications to exchange rate modeling, Journal of Applied Econometrics, 21(3), 307-336.

Alexander, C., & Lazar, E. [2009], Modelling Regime - Specific Stock Price Volatility, Oxford Bulletin of Economics and Statistics, 71(6), 761-797.

Alexander, C., Lazar, E., & Stanescu, S. [2011], Analytic moments for GARCH processes, ICMA Centre Discussion papers.

Alvarez, M., Luo, Y., Cahan, R., Jussa, J., Chen, Z., Wang, S. [2012]. Portfolios Under Construction: Risk and Alpha Alignment. Deutsche Bank Quantitative Strategy.

Anand, V., Natividade, C., Mesomeries, S., Davies, C., Jiang, S. and Capra, J [2014], Riding Carry, Deutsche Bank Quantcraft, 3 September 2014.

Ang, A. [2014], *Asset Management: A Systematic Approach to Factor Investing*, Oxford University Press, 1st Edition.

Asness, Clifford S., [2016], The Siren Song of Factor Timing, Journal of Portfolio Management, Vol. Special Issue, No. 1, 2016.

Bai, J., & Ng, S. [2002], Determining the number of factors in approximate factor models, Econometrica Vol. 70, No. 1, 191-221

Ball, C. A., & Torous, W. N. [1983], A simplified jump process for common stock returns, Journal of Financial and Quantitative analysis, 18(01), 53-65.

Bauwens, L., Omrane, W. B., & Giot, P. [2005], News announcements, market activity and volatility in the euro/dollar foreign exchange market, Journal of International Money and Finance, 24(7), 1108-1125.

Bauwens, L., Hafner, C. M., & Laurent, S. [2012], *Handbook of volatility models and their applications,* Volume 3,. John Wiley & Sons.

Baz, J., Granger, N., Harvey, C., Le Roux, N., Rattray, S. [2015], Dissecting Investment Strategies In The Cross Section And Time Series, available at SSRN.

Belsley, D. A., E. Kuh, & R. E. Welsh., *Regression Diagnostics*. New York, NY: John Wiley & Sons, Inc., 1980.

Benzoni, L. Collin-Dufreane, P. and Goldstein, R. [2010], Explaining asset pricing puzzles associated with the 1987 market crash, Working Paper, available at SSRN.

Bollerslev, T. [1986], Generalized autoregressive conditional heteroskedasticity, Journal of econometrics, 31(3), 307-327.

Bollerslev, T. [1990], Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model, The review of economics and statistics, 498-505.

Broadie, M., Chernov, M. and Johannes, M. [2009], Understanding Index Option Returns, Review of Financial Studies 22(11), 4493-4529.

Burmeister et al [1994] A Practitioner's Guide to Arbitrage Pricing Theory, The Research Foundation of the Institute of Chartered Financial Analysts.

Carr, P., Jin, X., & Madan, D. B. [2001]. Optimal investment in derivative securities, Finance and Stochastics, 5(1), 33-59.

Carr, P. and Wu, L. [2009], Variance Risk Premia, Review of Financial Studies 22, 1311-1341.

Cattell, R. B. [1966], The scree test for the number of factors, Multivariate Behavioral Research, 1, 245-276.

Chan, L., J. Karceski, and J. Lakonishok, [1999], On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model, Review of Financial Studies 12, 937974.

Chen, N., Roll, R. and Ross, S. [1986], Economic Forces and the Stock Market, Journal of Business, Vol. 59, No. 3 (1986), pp. 383-403.

Chen, L. [2010], Option Pricing And Hedging With Transaction Costs, PhD Disstertation – Stanford University.

Chen, L., Lai, T. L. and Lim, T. W. [2011], A New Approach to Pricing and Hedging Options with Transaction Costs, Available at SSRN.

Clenow, A. [2013], *Following the Trend: Diversified Managed Futures Trading*, John Wiley & Sons, 1st Edition.

Connor, G. [1996], The Three Types of Factor Models: A Comparison of Their Explanatory Power, Barra Research Insights.

Connor, G. [2000]. Robust confidence intervals for the bias test of risk forecasts. Technical report, MSCI Barra.

Cont, R. and Tankov, P. [2004], Financial Modelling with jump processes, Rama Cont, Peter Tankov, Boca Raton, Fla.: Chapman & Hall/CRC.

Corradi, V., Distaso, W., & Mele, A. [2006], Macroeconomic determinants of stock market volatility and volatility risk-premia, Working Paper, University of Warwick, UK.

Cosemans, M. [2011], The Pricing of Long and Short Run Variance and Correlation Risk in Stock Returns, Working paper, Erasmus University.

Davis, J. and Bhansali, V. [2014], Insuring Against Hidden Financial Catastrophe Risk, *Tail Risk Hedging: Theory and Practice*, Risk Books, 1st Edition.

Derman, E. and Miller, M. [2016], *The Volatility Smile*, John Wiley & Sons, 1st Edition.

DuMouchel, W. H., and F. L. O'Brien [1989], Integrating a Robust Option into a Multiple Regression Computing Environment, Computer Science and Statistics: Proceedings of the 21st Symposium on the Interface. Alexandria, VA: American Statistical Association, 1989.

Engle, R. F. [1982], Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica: Journal of the Econometric Society, 987-1007.

Engle, R. [2002], Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, Journal of Business & Economic Statistics, 20(3), 339-350.

Engle, R. F., & Rangel, J. G. [2005], The spline garch model for unconditional volatility and its global macroeconomic causes, Working Paper.

Engle, R. F., Ng, V. K., & Rothschild, M. [1990], Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills, Journal of Econometrics, 45(1-2), 213-237.

Franq, C. and Zakoian, J. [2010], *GARCH Models: Structure, Statistical Inference and Financial Applications*, John Wiley & Sons, 1st Edition.

Frost, P. A., and J. E. Savarino, [1986], An empirical Bayes approach to efficient portfolio selection, Journal of Financial and Quantitative Analysis 21, 293-305.

Gatheral, J. and Jacquier, A. [2013], Arbitrage-free SVI volatility surfaces, Working Paper, available at SSRN.

Glosten, L. R., Jagannathan, R., and Runkle, D. [1993], On the relation between the expected value and the volatility of the nominal excess return on stocks, The journal of finance 48.5 (1993): 1779-1801.

Greenberg et al [2016], Factors to Assets: Mapping Factor Exposures to Asset Allocations, Special Issue 2016.

Grinold, R. and Kahn, R. [1999], *Active Portfolio Management*, McGraw Hill, 2nd Edition.

Guegan, D., Ielpo, F., & Lalaharison, H. [2013], Option pricing with discrete time jump processes, Journal of Economic Dynamics and Control, 37(12), 2417-2445.

Haas, M., Mittnik, S., & Paolella, M. S. [2004a], A new approach to Markov-switching GARCH models, Journal of Financial Econometrics, 2(4), 493-530.

Haas, M., Mittnik, S., & Paolella, M. S. [2004b]. Mixed normal conditional heteroskedasticity. Journal of Financial Econometrics, 2(2), 211-250.

Hamilton, J. [2010], Regime switching models, In Macroeconometrics and Time Series Analysis (pp. 202-209), Palgrave Macmillan UK.

Hodges, S. D. and Neuberger, A. [1989], Optimal Replication of Contingent Claims under Transaction Costs, Review of Futures Markets, 8, 222–239.

Honore, P. [1998], Pitfalls in estimating jump-diffusion models, Centre for Analytical Finance, University of Aarhus, Aarhus School of Business, Working Paper Series No.18.

Ilmanen, A. [2011], *Expected Returns: An Investor's Guide to Harvesting Market Rewards*, John Wiley & Sons, 1st Edition.

James, W., and C. Stein, [1961], Estimation with quadratic loss, Proceedings of the Fourth Berkeley Symposium on Mathematics and Statistics 361-379.

Johnson, N. L. [1949], Systems of frequency curves generated by methods of translation. Biometrika, 36(1/2), 149-176.

Jorion, P., [1986], Bayes-Stein estimation for portfolio analysis, Journal of Financial and Quantitative Analysis 21, 279-292.

Kaminski, K. and Lo, A. [2007], When Do Stop-Loss Rules Stop Losses?, Working Paper, available at SSRN.

Leland, H. [1985], Option Pricing and Replication with Transaction Costs, Journal of Finance, 5, 1283–1301.

Luo, J., Saks, P. and Satchell, S. [2009], Implementing Risk Appetite in the Management of Currency Portfolio, Journal of Asset Management, 9, 380-397.

MacKinlay, A. C., and L. Pastor, [2000], Asset pricing models: implications for expected returns and portfolio selection, Review of Financial Studies 13, 883-916.

Mandelbrot [1963], The Variation of Certain Speculative Prices, The Journal of Business, 36(4), 394 – 419.

Marcucci, J. [2005], Forecasting stock market volatility with regime-switching GARCH models, Studies in Nonlinear dynamics and Econometrics, 9(4), 1-53.

Mccormick, S. and Natividade, C. [2008], "Volatility Drivers Part 1: Does Liquidity Affect Volatility?", DB FX Derivatives Focus, 26 February 2008.

Merton, R. C. [1976], Option pricing when underlying stock returns are discontinuous, Journal of financial economics, 3(1-2), 125-144.

Natividade, C. [2008], Volatility Drivers Part 2: The Cyclical Factors, DB FX Derivatives Focus, 15 April 2008.

Natividade, C. [2012a], Meet Vanessa, Deutsche Bank Quantcraft, 11 May 2012.

Natividade, C. [2012b], Catching Trends, Deutsche Bank Quantcraft, 25 October 2012.

Natividade, C. [2013a], Machine (un)learning?, Deutsche Bank Quantcraft, 6 March 2013.

Natividade, C. [2013b], Clock-watchers, Deutsche Bank Quantcraft, 10 April 2013.

Natividade, C. [2013c], Pairing Singles, Deutsche Bank Quantcraft, 13 May 2013.

Natividade, C. and Chen, K. [2011a], Breaking Out?, Deutsche Bank Quantcraft, 01 November 2011.

Natividade, C. and Chen, K. [2011c], Pairing Up, Deutsche Bank Quantcraft, 28 November 2011.

Natividade, C. and Chen, K. [2012a], Well Seasoned Vol Forecasting: SET Goes Global, Deutsche Bank Quantcraft, 04 January 2012.

Natividade, C. and Chen, K. [2012b], Switching Styles, Deutsche Bank Quantcraft, March 2012.

Natividade, C., Chen, K. and Wang, Y. [2011], Well Seasoned Vol Forecasting, Deutsche Bank Quantcraft, 14 November 2011.

Natividade, C., Mesomeris, S., Alvarez, M., Beceren, M. and Davis, C [2013], Colours of Trend, Deutsche Bank Quantcraft, 17 September 2013.

Natividade, C., Mesomeris, S., Davies, C., Jiang, C., Capra, J., Anand, V. and Qu, S. [2014], Value Investing: Tricky But Worthy, Deutsche Bank Quantcraft, 3 December 2014.

Natividade, C., Anand, V., Mesomeris, S., Davies, C., Ward, P., Capra, J., Qu, S. and Osiol, J. [2015], Delving Into New Territories, Deutsche Bank Quantcraft, 10 November 2015.

Natividade, C. and Anand, V. [2016], Trend Following: Asymmetries and Thresholds, Deutsche Bank Quant Snaps, 4 May 2016.

Natividade, C., Anand, V. and Mesomeris, C. [2016a], Carry: Think Different, Deutsche Bank Quant Snaps, 16 May 2016.

Oomen, R. and Sheppard, K [2014], Trading on-exchange versus over-the-counter: some observations from the equity and currency markets, Unpublished manuscript.

Prasad, R., Stanescu, S., Carter, S., and Jain, S. [2016], Designer Dispersion, Deutsche Bank Special Report, 5 February 2016.

Podkaminer E., [2013], Risk Factors as Building Blocks for Portfolio Diversification, Investment Risk and Performance, CFA Institute.

Ross, Stephen A., [1976], The Arbitrage Theory of Capital Asset Pricing, Journal of Economic Theory (December): 341-60.

Saravelos, G. and Grover, R. [2017], Why FX volatility is not cheap: understanding the drivers, DB FX Special Report, 14 March 2017.

Sepp, A. [2013], When You Hedge Discretely: Optimization of Sharpe Ratio for Delta-Hedging Strategy under Discrete Hedging and Transaction Costs, Journal of Investment Strategies, 2013, Vol. 3, No.1, pp.19-59.

Sharpe, W., [1963], A simplified model for portfolio analysis, Management Science 9, 277-293.

Simonatto, J. G. [2013], Approximating the Multivariate Distribution of Time-Aggregated Stock Returns Under GARCH, Journal of Risk, 16(2), 25-49.

Sinclair, E. [2008], *Volatility Trading*, John Wiley Y Sons, 1st Edition.

Tse, Y. K., & Tsui, A. K. C. [2002], A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations, Journal of Business & Economic Statistics, 20(3), 351-362.

Tuenter, H. [2001], An algorithm to determine the parameters of SU-curves in the johnson system of probabillity distributions by moment matching. Journal of Statistical Computation and Simulation, 70(4), 325-347.

Valenzuela, M. [2014], Implied Correlation and Expected Returns, Working Paper, available at SSRN.

Ward, P., Mesomeris, S., Davies, C., Natividade, C., Anand, V., Capra, J., Qu, S. and Osiol, J. [2016], Exploring The Risk Horizon, DB Quantitative Strategy: Portfolios Under Construction, 4 February 2016.

Whalley, A. E. and Wilmott, P. [1997]. An Asymptotic Analysis of an Optimal Hedging Model for Option Pricing with Transaction Costs, Mathematical Finance, 7 (3), 307–324.

Zakamouline, V. [2006a], Optimal Hedging of Option Portfolios with Transaction Costs, Working paper, Faculty of Economics, Agder University College, Norway.

Zakamouline, V. [2006b], Efficient Analytic Approximation of the Optimal Hedging Strategy for a European Call Option with Transaction Costs, Quantitative Finance 6:435–445.

Zakamouline, V. [2007], The Best Hedging Strategy in the Presence of Transaction Costs, Working paper, Faculty of Economics, Agder University College, Norway.

Zivot, E., [2011], Factor Models for Asset Returns, Lecture Handouts, University of Washington, 14 March 2011.

Appendix 1

Important Disclosures

*Other information available upon request

Prices are current as of the end of the previous trading session unless otherwise indicated and are sourced from local exchanges via Reuters, Bloomberg and other vendors . Other information is sourced from Deutsche Bank, subject companies, and other sources. For disclosures pertaining to recommendations or estimates made on securities other than the primary subject of this research, please see the most recently published company report or visit our global disclosure look-up page on our website at http://gm.db.com/ger/disclosure/DisclosureDirectory.eqsr. Aside from within this report, important conflict disclosures can also be found at https://gm.db.com/equities under the "Disclosures Lookup" and "Legal" tabs. Investors are strongly encouraged to review this information before investing.

Analyst Certification

The views expressed in this report accurately reflect the personal views of the undersigned lead analyst(s). In addition, the undersigned lead analyst(s) has not and will not receive any compensation for providing a specific recommendation or view in this report. Caio Natividade/Silvia Stanescu/Vivek Anand/Paul Ward/Simon Carter/Pam Finelli/Spyros Mesomeris

Hypothetical Disclaimer

Backtested, hypothetical or simulated performance results have inherent limitations. Unlike an actual performance record based on trading actual client portfolios, simulated results are achieved by means of the retroactive application of a backtested model itself designed with the benefit of hindsight. Taking into account historical events the backtesting of performance also differs from actual account performance because an actual investment strategy may be adjusted any time, for any reason, including a response to material, economic or market factors. The backtested performance includes hypothetical results that do not reflect the reinvestment of dividends and other earnings or the deduction of advisory fees, brokerage or other commissions, and any other expenses that a client would have paid or actually paid. No representation is made that any trading strategy or account will or is likely to achieve profits or losses similar to those shown. Alternative modeling techniques or assumptions might produce significantly different results and prove to be more appropriate. Past hypothetical backtest results are neither an indicator nor guarantee of future returns. Actual results will vary, perhaps materially, from the analysis.

Additional Information

The information and opinions in this report were prepared by Deutsche Bank AG or one of its affiliates (collectively "Deutsche Bank"). Though the information herein is believed to be reliable and has been obtained from public sources believed to be reliable, Deutsche Bank makes no representation as to its accuracy or completeness.

If you use the services of Deutsche Bank in connection with a purchase or sale of a security that is discussed in this report, or is included or discussed in another communication (oral or written) from a Deutsche Bank analyst, Deutsche Bank may act as principal for its own account or as agent for another person.

Deutsche Bank may consider this report in deciding to trade as principal. It may also engage in transactions, for its own account or with customers, in a manner inconsistent with the views taken in this research report. Others within Deutsche Bank, including strategists, sales staff and other analysts, may take views that are inconsistent with those taken in this research report. Deutsche Bank issues a variety of research products, including fundamental analysis, equity-linked analysis, quantitative analysis and trade ideas. Recommendations contained in one type of communication may differ from recommendations contained in others, whether as a result of differing time horizons, methodologies or otherwise. Deutsche Bank and/or its affiliates may also be holding debt or equity securities of the issuers it writes on. Analysts are paid in part based on the profitability of Deutsche Bank AG and its affiliates, which includes investment banking, trading and principal trading revenues.

Opinions, estimates and projections constitute the current judgment of the author as of the date of this report. They do not necessarily reflect the opinions of Deutsche Bank and are subject to change without notice. Deutsche Bank provides liquidity for buyers and sellers of securities issued by the companies it covers. Deutsche Bank research analysts sometimes have shorter-term trade ideas that are consistent or inconsistent with Deutsche Bank's existing longer term ratings. Trade ideas for equities can be found at the SOLAR link at http://gm.db.com. A SOLAR idea represents a high conviction belief by an analyst that a stock will outperform or underperform the market and/or sector delineated over a time frame of no less than two weeks. In addition to SOLAR ideas, the analysts named in this report may from time to time discuss with our clients, Deutsche Bank salespersons and Deutsche Bank traders, trading strategies or ideas that reference catalysts or events that may have a near-term or medium-term impact on the market price of the securities discussed in this report, which impact may be directionally counter to the analysts' current 12-month view of total return or investment return as described herein. Deutsche Bank has no obligation to update, modify or amend this report or to otherwise notify a recipient thereof if any opinion, forecast or estimate contained herein changes or subsequently becomes inaccurate. Coverage and the frequency of changes in market conditions and in both general and company specific economic prospects make it difficult to update research at defined intervals. Updates are at the sole discretion of the coverage analyst concerned or of the Research Department Management and as such the majority of reports are published at irregular intervals. This report is provided for informational purposes only and does not take into account the particular investment objectives, financial situations, or needs of individual clients. It is not an offer or a solicitation of an offer to buy or sell any financial instruments or to participate in any particular trading strategy. Target prices are inherently imprecise and a product of the analyst’s judgment. The financial instruments discussed in this report may not be suitable for all investors and investors must make their own informed investment decisions. Prices and availability of financial instruments are subject to change without notice and investment transactions can lead to losses as a result of price fluctuations and other factors. If a financial instrument is denominated in a currency other than an investor's currency, a change in exchange rates may adversely affect the investment. Past performance is not necessarily indicative of future results. Unless otherwise indicated, prices are current as of the end of the previous trading session, and are sourced from local exchanges via Reuters, Bloomberg and other vendors. Data is sourced from Deutsche Bank, subject companies, and in some cases, other parties.

The Deutsche Bank Research Department is independent of other business areas divisions of the Bank. Details regarding our organizational arrangements and information barriers we have to prevent and avoid conflicts of interest with respect to our research is available on our website under Disclaimer found on the Legal tab.

Macroeconomic fluctuations often account for most of the risks associated with exposures to instruments that promise to pay fixed or variable interest rates. For an investor who is long fixed rate instruments (thus receiving these cash flows), increases in interest rates naturally lift the discount factors applied to the expected cash flows and thus cause a loss. The longer the maturity of a certain cash flow and the higher the move in the discount factor, the higher will be the loss. Upside surprises in inflation, fiscal funding needs, and FX depreciation rates are among the most common adverse macroeconomic shocks to receivers. But counterparty exposure, issuer creditworthiness, client segmentation, regulation (including changes in assets holding limits for different types of investors), changes in tax policies, currency convertibility (which may constrain currency conversion, repatriation of profits and/or the liquidation of positions), and settlement issues related to local clearing houses are also important risk factors to be considered. The sensitivity of fixed income instruments to macroeconomic shocks may be mitigated by indexing the contracted cash flows to inflation, to FX depreciation, or to specified interest rates – these are common in emerging markets. It is important to note that the index fixings may -- by construction -- lag or mis-measure the actual move in the underlying variables they are intended to track. The choice of the proper fixing (or metric) is particularly important in swaps markets, where floating coupon rates (i.e., coupons indexed to a typically short-dated interest rate reference index) are exchanged for fixed coupons. It is also important to acknowledge that funding in a currency that differs from the currency in which coupons are denominated carries FX risk. Naturally, options on swaps (swaptions) also bear the risks typical to options in addition to the risks related to rates movements.

Derivative transactions involve numerous risks including, among others, market, counterparty default and illiquidity risk. The appropriateness or otherwise of these products for use by investors is dependent on the investors' own circumstances including their tax position, their regulatory environment and the nature of their other assets and liabilities, and as such, investors should take expert legal and financial advice before entering into any transaction similar to or inspired by the contents of this publication. The risk of loss in futures trading and options, foreign or domestic, can be substantial. As a result of the high degree of leverage obtainable in futures and options trading, losses may be incurred that are greater than the amount of funds initially deposited. Trading in options involves risk and is not suitable for all investors. Prior to buying or selling an option investors must review the "Characteristics and Risks of Standardized Options", at http://www.optionsclearing.com/about/publications/character-risks.jsp. If you are unable to access the website please contact your Deutsche Bank representative for a copy of this important document.

Participants in foreign exchange transactions may incur risks arising from several factors, including the following: (i) exchange rates can be volatile and are subject to large fluctuations; (ii) the value of currencies may be affected by numerous market factors, including world and national economic, political and regulatory events, events in equity and debt markets and changes in interest rates; and (iii) currencies may be subject to devaluation or government imposed exchange controls which could affect the value of the currency. Investors in securities such as ADRs, whose values are affected by the currency of an underlying security, effectively assume currency risk.
Unless governing law provides otherwise, all transactions should be executed through the Deutsche Bank entity in the investor's home jurisdiction. Aside from within this report, important conflict disclosures can also be found at https://gm.db.com/equities under the "Disclosures Lookup" and "Legal" tabs. Investors are strongly encouraged to review this information before investing.

United States: Approved and/or distributed by Deutsche Bank Securities Incorporated, a member of FINRA, NFA and SIPC. Analysts located outside of the United States are employed by non-US affiliates that are not subject to FINRA regulations.

Germany: Approved and/or distributed by Deutsche Bank AG, a joint stock corporation with limited liability incorporated in the Federal Republic of Germany with its principal office in Frankfurt am Main. Deutsche Bank AG is authorized under German Banking Law and is subject to supervision by the European Central Bank and by BaFin, Germany's Federal Financial Supervisory Authority.

United Kingdom: Approved and/or distributed by Deutsche Bank AG acting through its London Branch at Winchester House, 1 Great Winchester Street, London EC2N 2DB. Deutsche Bank AG in the United Kingdom is authorised by the Prudential Regulation Authority and is subject to limited regulation by the Prudential Regulation Authority and Financial Conduct Authority. Details about the extent of our authorisation and regulation are available on request.

Hong Kong: Distributed by Deutsche Bank AG, Hong Kong Branch.

India: Prepared by Deutsche Equities India Pvt Ltd, which is registered by the Securities and Exchange Board of India

(SEBI) as a stock broker. Research Analyst SEBI Registration Number is INH000001741. DEIPL may have received administrative warnings from the SEBI for breaches of Indian regulations.

Japan: Approved and/or distributed by Deutsche Securities Inc.(DSI). Registration number - Registered as a financial instruments dealer by the Head of the Kanto Local Finance Bureau (Kinsho) No. 117. Member of associations: JSDA, Type II Financial Instruments Firms Association and The Financial Futures Association of Japan. Commissions and risks involved in stock transactions - for stock transactions, we charge stock commissions and consumption tax by multiplying the transaction amount by the commission rate agreed with each customer. Stock transactions can lead to losses as a result of share price fluctuations and other factors. Transactions in foreign stocks can lead to additional losses stemming from foreign exchange fluctuations. We may also charge commissions and fees for certain categories of investment advice, products and services. Recommended investment strategies, products and services carry the risk of losses to principal and other losses as a result of changes in market and/or economic trends, and/or fluctuations in market value. Before deciding on the purchase of financial products and/or services, customers should carefully read the relevant disclosures, prospectuses and other documentation. "Moody's", "Standard & Poor's", and "Fitch" mentioned in this report are not registered credit rating agencies in Japan unless Japan or "Nippon" is specifically designated in the name of the entity. Reports on Japanese listed companies not written by analysts of DSI are written by Deutsche Bank Group's analysts with the coverage companies specified by DSI. Some of the foreign securities stated on this report are not disclosed according to the Financial Instruments and Exchange Law of Japan. Target prices set by Deutsche Bank's equity analysts are based on a 12-month forecast period.

Korea: Distributed by Deutsche Securities Korea Co.

South Africa: Deutsche Bank AG Johannesburg is incorporated in the Federal Republic of Germany (Branch Register Number in South Africa: 1998/003298/10).

Singapore: by Deutsche Bank AG, Singapore Branch or Deutsche Securities Asia Limited, Singapore Branch (One Raffles Quay #18-00 South Tower Singapore 048583, +65 6423 8001), which may be contacted in respect of any matters arising from, or in connection with, this report. Where this report is issued or promulgated in Singapore to a person who is not an accredited investor, expert investor or institutional investor (as defined in the applicable Singapore laws and regulations), they accept legal responsibility to such person for its contents.

Taiwan: Information on securities/investments that trade in Taiwan is for your reference only. Readers should independently evaluate investment risks and are solely responsible for their investment decisions. Deutsche Bank research may not be distributed to the Taiwan public media or quoted or used by the Taiwan public media without written consent. Information on securities/instruments that do not trade in Taiwan is for informational purposes only and is not to be construed as a recommendation to trade in such securities/instruments. Deutsche Securities Asia Limited, Taipei Branch may not execute transactions for clients in these securities/instruments.

Qatar: Deutsche Bank AG in the Qatar Financial Centre (registered no. 00032) is regulated by the Qatar Financial Centre Regulatory Authority. Deutsche Bank AG - QFC Branch may only undertake the financial services activities that fall within the scope of its existing QFCRA license. Principal place of business in the QFC: Qatar Financial Centre, Tower, West Bay, Level 5, PO Box 14928, Doha, Qatar. This information has been distributed by Deutsche Bank AG. Related financial products or services are only available to Business Customers, as defined by the Qatar Financial Centre Regulatory Authority.

Russia: This information, interpretation and opinions submitted herein are not in the context of, and do not constitute, any appraisal or evaluation activity requiring a license in the Russian Federation.

Kingdom of Saudi Arabia: Deutsche Securities Saudi Arabia LLC Company, (registered no. 07073-37) is regulated by the Capital Market Authority. Deutsche Securities Saudi Arabia may only undertake the financial services activities that fall within the scope of its existing CMA license. Principal place of business in Saudi Arabia: King Fahad Road, Al Olaya District, P.O. Box 301809, Faisaliah Tower - 17th Floor, 11372 Riyadh, Saudi Arabia.

United Arab Emirates: Deutsche Bank AG in the Dubai International Financial Centre (registered no. 00045) is regulated by the Dubai Financial Services Authority. Deutsche Bank AG - DIFC Branch may only undertake the financial services activities that fall within the scope of its existing DFSA license. Principal place of business in the DIFC: Dubai International Financial Centre, The Gate Village, Building 5, PO Box 504902, Dubai, U.A.E. This information has been distributed by Deutsche Bank AG. Related financial products or services are only available to Professional Clients, as defined by the Dubai Financial Services Authority.

Australia: Retail clients should obtain a copy of a Product Disclosure Statement (PDS) relating to any financial product referred to in this report and consider the PDS before making any decision about whether to acquire the product. Please refer to Australian specific research disclosures and related information at https://australia.db.com/australia/content/research-information.html

Australia and New Zealand: This research is intended only for "wholesale clients" within the meaning of the Australian Corporations Act and New Zealand Financial Advisors Act respectively.

Additional information relative to securities, other financial products or issuers discussed in this report is available upon request. This report may not be reproduced, distributed or published without Deutsche Bank's prior written consent.

David Folkerts-Landau
Group Chief Economist and Global Head of Research

Raj Hindocha
Global Chief Operating Officer
Research

Michael Spencer
Head of APAC Research
Global Head of Economics

Steve Pollard
Head of Americas Research
Global Head of Equity Research

Anthony Klarman
Global Head of
Debt Research

Paul Reynolds
Head of EMEA
Equity Research

Dave Clark
Head of APAC
Equity Research

Pam Finelli
Global Head of
Equity Derivatives Research

Andreas Neubauer
Head of Research - Germany

Stuart Kirk
Head of Thematic Research

International Locations

Deutsche Bank AG
Deutsche Bank Place
Level 16
Corner of Hunter & Phillip Streets
Sydney, NSW 2000
Australia
Tel: (61) 2 8258 1234

Deutsche Bank AG
Große Gallusstraße 10-14
60272 Frankfurt am Main
Germany
Tel: (49) 69 910 00

Deutsche Bank AG
Filiale Hongkong
International Commerce Centre,
1 Austin Road West,Kowloon,
Hong Kong
Tel: (852) 2203 8888

Deutsche Securities Inc.
2-11-1 Nagatacho
Sanno Park Tower
Chiyoda-ku, Tokyo 100-6171
Japan
Tel: (81) 3 5156 6770

Deutsche Bank AG London
1 Great Winchester Street
London EC2N 2EQ
United Kingdom
Tel: (44) 20 7545 8000

Deutsche Bank Securities Inc.
60 Wall Street
New York, NY 10005
United States of America
Tel: (1) 212 250 2500