DISCUSSION PAPER SERIES

DP12687

RISK EVERYWHERE: MODELING AND MANAGING VOLATILITY

Tim Bollerslev, Benjamin Hood, John Huss and Lasse Heje Pedersen

FINANCIAL ECONOMICS

ISSN 0265-8003

RISK EVERYWHERE: MODELING AND MANAGING VOLATILITY

Tim Bollerslev, Benjamin Hood, John Huss and Lasse Heje Pedersen

Discussion Paper DP12687
Published 05 February 2018
Submitted 05 February 2018

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

RISK EVERYWHERE: MODELING AND MANAGING VOLATILITY

Abstract

Based on a unique high-frequency dataset for more than fifty commodities, currencies, equity indices, and fixed income instruments spanning more than two decades, we document strong similarities in realized volatilities patterns across assets and asset classes. Exploiting these similarities within and across asset classes in panel-based estimation of new realized volatility models results in superior out-of-sample risk forecasts, compared to forecasts from existing models and more conventional procedures that do not incorporate the information in the high-frequency intraday data and/or the similarities in the volatilities. A utility-based framework designed to evaluate the economic gains from risk modeling highlights the interplay between parsimony of model specification, transaction costs, and speed of trading in the practical implementation of the different risk models.

JEL Classification: C22, C51, C53, C58

Keywords: Market and volatility risk, high-frequency data, realized volatility, risk modeling and forecasting, volatility trading, risk targeting, realized utility

Tim Bollerslev - boller@econ.duke.edu
Duke University - Finance; Duke University - Department of Economics; National Bureau of Economic Research (NBER)

Benjamin Hood - benjamin.hood@aqr.com
AQR Capital Management, LLC

John Huss - john.huss@gmail.com
AQR Capital Management, LLC

Lasse Heje Pedersen - lhp001@gmail.com
Copenhagen Business School, AQR, and CEPR

Acknowledgements
We thank participants at the 2016 Volatility Institute Conference at NYU, Stern, Aaron Brown, Eric Ghysels, Antti Ilmanen, Ari Levine, Yao Hua Ooi, Scott Richardson, and Liuren Wu for useful comments and discussions. We are also most grateful to Andrew Karolyi (the Editor) and two anonymous referees for their detailed comments, which greatly improved the paper. Bollerslev is at Duke University, NBER, and CREATES, Social Sciences Building, Durham, NC 27708. He also has a consulting relationship with AQR Capital Management. Hood and Huss are at AQR Capital Management, 2 Greenwich Plaza, Greenwich, CT 06830. Pedersen is at AQR Capital Management, Copenhagen Business School, NYU, and CEPR. and gratefully acknowledges support from the European Research Council (ERC grant No. 312417) and the FRIC Center for Financial Frictions (grant No. DNRF102).

Risk Everywhere: Modeling and Managing Volatility

This version: March 21, 2017

Tim Bollerslev[a], Benjamin Hood[b], John Huss[b], Lasse Heje Pedersen[b,c]

[a] *Duke University, NBER, and CREATES*
[b] *AQR Capital Management*
[c] *Copenhagen Business School, NYU, and CEPR*

Abstract

Based on a unique high-frequency dataset for more than fifty commodities, currencies, equity indices, and fixed income instruments spanning more than two decades, we document strong similarities in realized volatilities patterns across assets and asset classes. Exploiting these similarities within and across asset classes in panel-based estimation of new realized volatility models results in superior out-of-sample risk forecasts, compared to forecasts from existing models and more conventional procedures that do not incorporate the information in the high-frequency intraday data and/or the similarities in the volatilities. A utility-based framework designed to evaluate the economic gains from risk modeling highlights the interplay between parsimony of model specification, transaction costs, and speed of trading in the practical implementation of the different risk models.

Keywords: Market and volatility risk; high-frequency data; realized volatility; risk modeling and forecasting; volatility trading; risk targeting; realized utility. *JEL:* C22, C51, C53, C58

☆We thank participants at the 2016 Volatility Institute Conference at NYU, Stern, Aaron Brown, Eric Ghysels, Antti Ilmanen, Ari Levine, Yao Hua Ooi, Scott Richardson, and Liuren Wu for useful comments and discussions. We are also most grateful to Andrew Karolyi (the Editor) and two anonymous referees for their detailed comments, which greatly improved the paper. Bollerslev is at Duke University, NBER, and CREATES, Social Sciences Building, Durham, NC 27708. He also has a consulting relationship with AQR Capital Management. Hood and Huss are at AQR Capital Management, 2 Greenwich Plaza, Greenwich, CT 06830. Pedersen is at AQR Capital Management, Copenhagen Business School, NYU, and CEPR. and gratefully acknowledges support from the European Research Council (ERC grant No. 312417) and the FRIC Center for Financial Frictions (grant No. DNRF102).

1. Introduction

Measuring, forecasting, and controlling risk are at the heart of financial economic theory and practice. Investors and asset managers generally seek to maximize return while limiting risk. Many traders also have specific risk limits or risk targets. Traders in a bank, for example, often face risk limits, while hedge funds often tell their investors that a particular fund is expected to realize a certain risk level. So-called risk parity strategies are explicitly designed to equate the risk stemming from the different investments included in a portfolio, rather than the capital allocated to the different components. Such investors therefore continually monitor global risk as it evolves and change their portfolios in response, reducing notional exposures as risk rises and increasing positions as risk fades.

This paper presents a framework for measuring, modelling, and forecasting risk across global assets and asset classes. Our results are based on on a broad dataset of realized volatilities constructed from high-frequency data for more than fifty commodities, currencies, equity indices, and fixed income futures. We find that our new risk models, combined with panel-based estimation techniques designed to exploit the strong commonalities observed in the volatilities across assets and asset classes, result in statistically significant out-of-sample forecast improvements and non-trivial utility gains compared to more conventional individually estimated asset specific risk-models.

Our analysis is based on a comprehensive database comprised of high-frequency intraday data from different global market spanning more than two decades, covering 20 commodities, 21 equity indices, 8 fixed income futures, and 9 currencies. We begin by computing the realized volatilities (RV) for each day and asset in our sample. When qualitatively comparing the estimated RVs, the differences in risk levels across assets and asset classes immediately stand out. However, when we normalize each asset's daily realized volatilities by their respective sample averages, striking similarities emerge. Indeed, these "normalized risk measures" have almost identical unconditional distributions and similar highly persistent autocorrelation structures when comparing across assets and asset classes. Hence, the volatilities of different assets — whether equities, bonds, commodities, or currencies — appear to behave almost the same over time. Going one step further, we also document strong volatility spillover effects both within and across different markets and geographical regions. The existence of spillover effects and commonalities in the dynamic dependencies is, of course, well know from the already existing volatility literature and the estimation results obtained with traditional GARCH and stochastic volatility models; see, e.g., Taylor (2005), Andersen, Bollerslev, Christoffersen, and Diebold (2006), and references therein.

We next build risk forecasting models explicitly designed to exploit these strong similarities in the distributions of the volatilities across and within asset classes. The formulation of our models are motivated by the heterogeneous autoregressive (HAR) model of Corsi (2009) and draws on important insights from the mixed data sampling (MIDAS) approach of Ghysels, Santa-Clara, and Valkanov (2006) and Ghysels, Sinko, and Valkanov (2007). First, we show how to simultaneously estimate risk models across *many* assets using panel regressions that add power by exploiting the similarities in the cross-asset risk characteristics.[1] An important step is to "center" the models and eliminate the asset specific intercept terms to ensure that all parameters are "scale free" in the sense that they do not depend on the level of risk. Second, we introduce new "smooth" realized volatility models, in which the forecasted future volatilities depends on the past volatilities in a way that is continuous and decreasing in the lag lengths, thereby eliminating non-monotonicities arising from estimation noise and predictable jumps in the risk forecast as time passes. Our preferred specification, which we denote the heterogenous exponential realized volatility model (HExp for short), in particular, is based on a simple mixture of Exponentially Weighted Moving Average (EWMA) factors. Third, to account for the volatility spillover effects and strong commonalities observed not just across different assets, but also across different geographical regions, we augment the asset specific HExp model with a lagged "global" risk factor.

Looking at the in-sample results, we find that all of the RV models that we consider perform well compared to models that "only" use daily returns. By construction, when looking in-sample, the models that are tailored to each asset separately have larger predictive power in terms of R^2 than models that enforce a common risk model across assets. However, when looking at out-of-sample predictability, the models that impose common parameters generally perform better. In particular, enforcing common parameters across models within each asset class produces higher average out-of-sample R^2s than individually estimated models. Even more surprisingly, enforcing common parameters not just within but across *all* asset classes, the properly "centered" risk models result in even higher average out-of-sample R^2s. The basic HExp model and the HExp model with the "global" risk factor result in the highest average out-of-sample predictability among all of the models, suggesting that the commonality and "smoothness" embedded in the HExp formulations ensure a robustness beyond that of the standard existing risk models.

[1]The use of panel-based estimation technique to enhance the efficiency of the individual forecasts parallels the use of Bayesian estimation procedures more generally, as exemplified by Karolyi (1993) who rely on Bayes-Stein shrinkage for improving the forecasts of individual stock return volatilities based on the cross-sectional dispersion.

Last, but not least, we present a simple framework for quantifying the utility benefits of risk modeling. Our approach is linked to the literature that seeks to assess the utility benefits of return predictability in the presence of empirically realistic transaction costs and other practical implementation issues; see, e.g., Balduzzi and Lynch (1999), Sangvinatsos and Wachter (2005) and Lynch and Tan (2010). It is also related to the work of Fleming, Kirby, and Ostdiek (2001, 2003), and the idea of using a quadratic utility-function to evaluate the benefits of volatility timing. In contrast to all of these approaches, however, which depend explicitly on forecasts of – and/or realizations of – both the future returns and volatilities, our method focus exclusively on volatility forecasting. Specifically, by considering an investor with mean-variance preferences that trades an asset with a constant Sharpe ratio, the investor's optimal portfolio naturally adjusts the position size to keep a constant volatility, reflective of the investor's risk aversion. Correspondingly, the investor's utility is directly related to the volatility: the investor achieves the maximum utility by successfully targeting a constant risk level, while the utility decreases with the volatility-of-volatility. Hence, risk models that help the investor achieve more accurate volatility forecasts are associated with higher levels of utility.

We show that, in this situation, under conservative assumptions about the Sharpe ratio and the investor's risk aversion, having an RV-based risk model is worth about 0.55% per year relative to using the best possible static risk model.[2] Put differently, the assumed investor would in principle pay 0.55% of her/his wealth each year to have access to one of the dynamic RV-based risk models developed here. These benefits are even greater when we take realistic transaction costs into account, as the new "smooth" risk models not only produce more accurate risk forecasts, but also more stable forecasts resulting in less spurious trading.

2. Realized Volatilities: Data Sources and Construction

There is a long history in finance of heuristically quantifying the ex-post volatility based on the sum intra-period squared returns.[3] This approach may be formally justified by the theory of quadratic variation and the notion of ever finer sampled returns over fixed time intervals, or so-called in-fill asymptotic arguments; see, e.g., the discussion and references in Andersen, Bollerslev, Christoffersen, and Diebold (2013).

[2]This utility benefit is based on an investor with a risk aversion of 3 trading an asset with a Sharpe ratio of 0.3. Even greater benefits obtain for investors with lower levels of risk aversion and/or higher Sharpe ratios.

[3]French, Schwert, and Stambaugh (1987) and Schwert (1989), for instance, rely on the sum of daily squared returns in their construction and modeling of monthly U.S. equity volatilities, while Hsieh (1991) estimates models for daily volatilities constructed from 15-minute S&P 500 returns.

2.1. *Realized Volatilities and Quadratic Variation*

To formally lay out the basic idea underlying the realized volatility concept, let the unit time interval correspond to a day. The realized variation defined by the summation of high-frequency intraday squared returns,

$$RV_t \equiv \sum_{i=1}^{1/\Delta} \left[log(P_{t-1+i\Delta}) - log(P_{t-1+(i-1)\Delta}) \right]^2 , \tag{1}$$

then consistently estimates the quadratic variation, i.e., the true variation, on day t as the number of intraday observations increases ($1/\Delta \to \infty$), or equivalently the length of the intraday return interval decreases ($\Delta \to 0$). This effectively renders the daily variation directly observable on an ex-post basis.[4] The volatility over longer, say weekly or monthly, horizons may similarly be estimated by summing the intraday squared returns over a week or a month, or equivalently by summing the daily realized volatilities RV_t over the relevant longer multi-day horizons.

2.2. *Data Sources and "Cleaning"*

Our data covers a total of 58 different assets across commodities (20), equities (21), fixed income (8), and foreign exchange (9). The asset universe is comprised of global equity index futures (both developed and emerging markets), global developed fixed income futures, commodity futures, and spot market foreign exchange rates. Our specific choice of assets is dictated by liquidity concerns and correspondingly the availability of reliable high-frequency intraday prices.

Our primary source of data for equities, fixed income, and commodities is the Thomson Reuters Tick History (TRTH) database. To extend the history for some of the assets, most notably fixed income and commodities, we also use data from TickData.com (TDC). For the foreign exchange data, we rely exclusive on Olsen Data (OD). The data for all of the assets runs through September 2014. The start of the sample period differs across assets, with some starting as early as October 1992. In general, the data for commodities are available the earliest, followed by equities and fixed income, with foreign exchange having the shortest time span. The exact start dates for all of the assets and the relevant data sources are summarized in Table A.1 in the Appendix.[5] Multiple futures contracts for the same underlying asset, but

[4]Consistent with the common use of the two terms in the extant literature, we will interchangeably refer to RV_t as the daily realized variation, or the daily realized volatility, while also sometimes referring explicitly to $RV_t^{1/2}$ as the daily realized volatility.

[5]The staggered start dates are easily accommodated by our panel-based estimation procedures discussed below, which do not restrict the sample sizes to be the same for all assets.

with different expiration, trade at the same time, and new contracts are opened as others expire. We focus on the most liquid contract, "rolling" from one contract to another at a regular schedule as further discussed in the Appendix. We organize the resulting single time series of returns for each asset into minute bars based on the last observation prior to the end of each minute. For the TRTH and OD databases, we use the mid-quote price (average between the bid and ask price). For the TDC data we use the observed trade price (TDC does not provide quote-level data prior to 2010). To avoid "polluting" the high-frequency data with quote changes that occur during illiquid periods, we only use the minute bars for which there are at least one valid trade within that minute. Lastly, having organized all of the data into minute bars, we apply a series of "sanity filters" to clean out any obvious data errors. These filters are further discussed in the Appendix.

2.3. Overnight Returns and Intraday Sampling

It is well established that volatility tend to be higher during exchange trading hours than during non-trading hours; see, e.g., French and Roll (1986). The theory underlying the consistency of the realized volatility measure portrays prices as evolving continuously through time. In actuality, of course, most markets close on weekends and certain holidays, change their trading hours, and sometimes experience "ghost" hours, where liquidity is very poor despite the markets technically being open. Accordingly, we only retain the trading hours for which the liquidity is sufficiently high to ensure a reasonable quality of the high-frequency data. We use the Financial Calendars (FinCal) database for market open and close times, together with so-called "liquidity plots" to delineate the periods of actively operating markets; further details are provided in the Appendix. Having determined the period for which reliable high-frequency data is available, we simply add the corresponding "overnight" squared returns to the daily realized volatilities constructed from the "intraday" squared returns to obtain an RV measure for the whole day.[6]

We rely on a common 5-minute sampling frequency for calculating the intraday RVs for all of the assets. This choice directly mirrors the sampling frequency used in much of the existing realized volatility literature. It may be justified by the volatility signature plots (Andersen, Bollerslev, Diebold, and Labys, 2000) discussed in the Appendix. To further enhance the efficiency of the RV estimates, we average the five different daily RVs obtained by

[6]Following Hansen and Lunde (2005b), we also experimented with "optimally" combining the squared overnight returns and the intraday RV, proportionally scaling the intraday RV, and separately modeling the overnight squared returns. None of these more complicated to implement procedures clearly dominated the much easier-to-implement approach of simply adding the overnight squared returns to the intraday RV's.

starting the day at first five unique one-minute marks. A number of other consistent realized volatility estimators, requiring the choice of additional tuning and/or nuisance parameters, have been proposed in the literature.[7] However, the theoretical comparisons in Andersen, Bollerslev, and Meddahi (2011) and Ghysels and Sinko (2011) show that from a theoretical forecasting perspective, the simple sub-sampled five-minute RV estimator that we rely on here performs on par with or better than all of these more complicated estimators. The empirical study by Liu, Patton, and Sheppard (2015), comparing more than four-hundred different RV estimators across multiple assets, similarly concludes that: "it is difficult to significantly beat 5-minute RV."

3. Risk Characteristics Everywhere

To help guide the specification of empirically realistic risk models, we want to understand the distributional characteristics of the risks both within each of the four asset classes, equities, bonds, commodities, and currencies, as well as the similarities and differences across asset classes.

3.1. Unconditional Distributions and Dynamic Dependencies

To begin, Figure 1 shows the time series of annualized realized volatilities for four representative assets, one from each asset class: S&P 500, 10-year T-bonds, Crude Oil, and Dollar/Euro. Even though the four volatilities obviously exhibit their own distinct behaviors, there is a clear commonality in the dynamic patterns observed across the four assets, with most of the peaks readily associated with specific economic events.

In spite of the similarities in the general patterns, the overall levels of the volatilities clearly differ across the four different assets. This is further evidenced by Figure 2, which plots the unconditional distribution of the daily realized volatilities for the same four representative assets. As the figure shows, Crude Oil is the most volatile on average, followed by the S&P 500, and the Dollar/Euro exchange rate. The volatility of 10-year T-bonds is by far the lowest.

This same ranking caries over to the four asset classes more generally. In particular, looking at the summary statistics for the daily realized volatilities averaged across each of the assets within each of the four asset classes reported in Table 1, the average annualized

[7] These includes the two-scale RV of Zhang, Mykland, and Aït-Sahalia (2005), the kernel-based RV of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), and the pre-averaged RV of Jacod, Li, Mykland, Podolskij, and Vetter (2009).

volatility for commodities and equities equal 25.4% and 20.6%, respectively, compared to 10.3% for foreign exchange, and just 3.1% for fixed income.

These differences in the mean levels of the volatilities across the different assets and asset classes are, of course, well know. More interesting features arise when we consider the volatilities normalized by their sample mean, $RV_t/Mean(RV_t)$, corresponding to the risk of a leveraged (or de-leveraged) position. For example, if $Mean(RV_t) = 0.5$, then the normalized volatility measure corresponds to the risk of a position that is always leveraged 2-to-1. Hence, normalizing each individual contract by its own average volatility is equivalent to measuring risk on a common scale, in the sense that each position is leveraged to the same common average risk level.

Interestingly, the unconditional distributions of these daily normalized realized volatilities are remarkably similar, both across assets and asset classes. Specifically, looking at Figure 3 and the unconditional distributions of the normalized volatilities for each of the four assets classes, the sampling distributions are obviously very close.[8] The unconditional distributions for the normalized volatility of each of the individual assets within each of the four asset classes are also very similar, as reported in the Supplementary Appendix. It is important to note that the "width" of these distributions are not similar by construction, as would be the case if we normalized by subtracted the mean and divided by the standard deviation, $[RV_t - Mean(RV_t)]/std(RV_t)$, which would match both the mean and standard deviation by construction (or, said differently, such a normalization implies a loss of two degrees of freedom, rather than just one for our normalization). Hence, the fact that the $RV_t/Mean(RV_t)$ normalized distributions are so similar is not hard-wired, but rather direct evidence that risks do indeed behave similarly across asset classes. Importantly, these similarities also imply that risk parity strategies designed to match the average volatility of different assets and/or asset classes will not only equate the average volatility levels, but effectively the entire unconditional distributions of the ex-post realized volatilities for the leveraged positions.

These commonalities in the unconditional distributions carry over to the general dynamic dependencies. Looking at the autocorrelations for the daily realized volatilities averaged across the different assets within each of the four asset classes shown in Table 1 and Figure 4, the general patterns and decay rates are very similar, again with equities perhaps being slightly different for the short-term lags.[9] Similar dynamic dependencies have, of course,

[8]Although seemingly very close, pairwise Kolmogorov–Smirnov tests easily reject that the four sampling distributions in Figure 3 are identical, with a p value of 0.0087 for Commodities and Fixed Income, and p-values below 0.0001 for all of the other pairwise comparison.

[9]Because of the varying degrees of measurement errors in the realized volatilities for the assets within the

been extensively documented in the burgeon volatility literature.[10]

3.2. Cross-Asset Dependencies, Spillovers, and Global Volatility

The averages of the standard sample correlations for the realized volatilities reported in the top panel in Table 2 are all positive. This comovement of risk across assets and asset classes is consistent with the visual impression from the time series plots for the four representative assets previously discussed in Figure 1. Looking at the actual numbers, commodity volatilities are generally the least correlated, both within the asset class and across other asset classes. That is, the risk of different commodities comove less with each other than equity risks comove with each other, and so on, and commodity risks also comove less with the risk of equity, fixed income, and currencies than these comove with each other. In fact, commodity risk comoves about as much across asset classes as within the asset class.

In addition to the within and across asses class correlations, the last row reports the average correlations with a "global risk factor" that we denote $GlRV$. The global risk factor is defined as the average normalized RVs across all assets. Since we will also $GlRV$ for forecasting, we construct this factor on an asset-specific basis in the following way that prevents look-ahead bias due to time-zone effects. In particular, for any specific asset, we construct the corresponding $GlRV$ so that today's $GlRV$ does not use any data that overlaps with tomorrow's trading hours for the specific asset, lagging by one day any asset that would otherwise create such an overlap.[11] Based on this lagging convention, on each day and for each asset i, we compute the asset-specific $GlRV$ as the average normalized RV scaled back to the asset's own level of volatility; i.e., $\left(\frac{1}{J}\sum_{j=1,\ldots,J}\frac{RV_{t,j}}{\overline{RV_j}}\right)\overline{RV_i}$.

This global volatility factor captures well the overall volatility dynamic across asset classes as seen in the last row of the top panel in Table 2. Indeed, as a sign of the strong commonalities in the realized volatilities, the average correlations with this new global risk factor systematically exceed the across asset class correlations. With the exception of foreign ex-

four asset classes, the levels of the autocorrelations for the shortest lags are not directly comparable; see also the discussion in Hansen and Lunde (2014).

[10]The book by Taylor (1986), in particular, provides some of the earliest empirical evidence explicitly highlighting the similarities in the autocorrelations of absolute and squared daily returns across assets and asset classes. GARCH and stochastic volatility models also typically result in very similar dynamic parameter estimates for different assets. For instance, applying a GARCH(1,1) model to daily returns, the estimates for α and β in the usual notation of the model are typically around $0.03-0.08$ and $0.92-0.97$, respectively, with the sum very close to unity; see, e.g., Hansen and Lunde (2005a).

[11]For instance, the day $t-1$ market hours for U.S. equities overlaps with the overnight portion of day t for Australia equities. Hence, in the construction of the global factor used for Australian equities, we shift the RV for U.S. equities back by one day. The RV for Australian equities, on the other hand, need not be shifted in the construction of the global factor for U.S. equities. To also account for changes in market hours over the sample period, this shifting is done on a day-by-day basis for each asset.

change, the global risk factor correlations are also all higher than the within asset class correlations.

The literature is rife with studies seeking to model these strong dependencies and possible volatility spillover effects using multivariate GARCH and related procedures; see, e.g., Engle, Ito, and Lin (1990) and Karolyi (1995) for some of the earliest evidence. In contrast to these earlier studies, which rely on parametric volatility models for inferring the dependencies, our use of realized volatilities allow us to directly quantify the strengths of any spillover effects. To this end, the middle panel in Table 2 reports the average partial correlation coefficients obtained from regressions of the daily RVs for each of the assets within the asset classes indicated in the columns on their own daily lagged value and the lagged values of the RVs for the assets in the asset classes indicated in the rows. To allow for a scale-invariant interpretation, we further normalize the RVs by their sample means; i.e., we report the average of the estimated $b_{2,ij}$ coefficients from the regressions $RV_{t,j}/\overline{RV_j} = b_{0,ij} + b_{1,ij} RV_{t-1,j}/\overline{RV_j} + b_{2,ij} RV_{t-1,i}/\overline{RV_i} + u_{t,ij}$. The bottom panel shows the resulting average increases in the R^2s compared to simple first order autoregressions that only control for the own lagged dynamic dependencies; i.e., $RV_{t,j}/\overline{RV_j} = b_{0,j} + b_{1,j} RV_{t-1,j}/\overline{RV_j} + u_{t,j}$.

Consistent with the presence of strong cross-market linkages and spillover effects, all of the average partial correlations are positive.[12] Comparing the results across the different asset classes, equity volatilities as a whole tend to exert the largest impact on the other asset class volatilities. Meanwhile, the magnitude of the equity partial correlations and the resulting increases in the R^2s are all dominated by those of the global risk factor. This naturally raises the questions of what is behind these linkages, and, in particular, what might explain the dynamic variation in the global risk factor?

The economic forces behind volatility clustering per se remain poorly understood, and a full fledged analysis of that question is also beyond the scope of the present paper. However, in an effort to shed some light on the mechanisms at work, Table 3 reports the monthly (end-of-month) correlations between an exponentially smoothed version of the global volatility factor $ExpGlRV$, as used in our preferred HExpGl model formally defined in Section 4.6 below, and four other variables naturally related to volatility. The global volatility factor together with three of the four four variables are also plotted in Figure 5.

The first entry in Table 3, in particular, shows that the global volatility factor is negatively correlated with the U.S. investor sentiment index of Baker and Wurgler (2006).[13]

[12] As reported in the Supplementary Appendix, the vast majority of the individually estimated coefficients are also statistically significant.

[13] We rely on the "orthogonalized" sentiment measure available on Wurgler's website:

Baker, Wurgler, and Yuan (2012) have previously argued that U.S. investor sentiment is contagious, and that international capital flows may in part account for this contagion. The well established strong link between trading volume and return volatility (see, e.g., Karpoff, 1987, for a survey of some of the earliest empirical evidence) may help explain the connection between investor sentiment and global volatility. Further along these lines, Karolyi, Lee, and van Dijk (2012) have previously documented strong commonalities in equity trading volumes across countries, arguably driven by correlated trading activity among institutional investors and common investor sentiment.

The previous correlation portrays a monotone relationship between volatility and sentiment, possibly driven by correlated trading. However, if "noise" traders acting on sentiment affect prices, unusually high or low levels of sentiment should both be associated with high levels of volatility; see, e.g., Brown (1999). Consistent with this idea, the monthly correlation between the global volatility factor and "unusual" sentiment, defined as the absolute value of the same sentiment measure minus its sample mean, equals 0.197.[14] Thus, whereas the general level of U.S. investor sentiment is negatively correlated with global market volatility, unusual U.S. investor sentiment is positively correlated with global market volatility.

The variance risk premium, formally defined as the difference between the actual and risk-neutral expectation of the future return variation, is naturally interpreted as a measure of aggregate risk aversion; see, e.g., Bakshi and Madan (2006). Supporting the idea that risk aversion, and in turn risk bearing capacity, influence volatility, the global volatility factor is strongly negatively correlated with the U.S. equity variance risk premium.[15]

Numerous studies have sought to relate low-frequency variation in return volatility to directly observable macroeconomic variables or indicators; see, e.g., Schwert (1989) and Engle and Rangel (2008). The reported relations, however, are weak at best. At the other end of spectrum, a number of studies have documented a sharp, but short lived, increase in intraday volatility following macroeconomic and other public news announcements; see, e.g., Ederington and Lee (1993) and Andersen, Bollerslev, Diebold, and Vega (2003b). The last entry in Table 3 reports the correlation between the global volatility factor and a news surprise variable, constructed as the average of the standardized absolute surprises for five of the most important U.S. macroeconomic news announcements released over the month.[16]

http://people.stern.nyu.edu/jwurgler/.

[14]Restricting the sentiment measure to just two states, Yu and Yuan (2011) also find that especially high sentiment tend to be associated with high volatility.

[15]We follow Bollerslev, Tauchen, and Zhou (2009) in quantifying the variance risk premium as the difference between the VIX and the realize U.S. equity volatility over the past month.

[16]We follow Andersen, Bollerslev, Diebold, and Vega (2003b) in identifying the five most important reg-

Corroborating the idea that the volatility in global financial markets may in part be attributed to "news" about the U.S. economy, the correlation between this news surprise variable and the global volatility factor equals 0.183.

Taken as a whole, the results in Table 3 support the idea that commonalities in trading behavior possibly induced by changes in investor sentiment and/or notions of risk aversion together with unanticipated news, serve important roles in accounting for the strong commonalities in the dynamic dependencies observed in the volatilities within and across asset classes. We will not pursue this any further here. Instead, we turn next to a discussion of the practical risk models that we use for modeling and forecasting these dynamic dependencies.

4. Risk Modeling

Our new risk models are explicitly designed to incorporate the strong similarities observed in risk characteristics across assets and asset classes.

4.1. Omnibus RV-Based Risk Models: AR(∞)

To facilitate the discussion, it is instructive to consider the omnibus AR(∞) risk model,

$$RV_{t+1} = b_0 + b_1 RV_t + b_2 RV_{t-1} + ... + \epsilon_t = b_0 + b(L)RV_t + \epsilon_t, \tag{2}$$

in which the realized volatility on day $t+1$ is determined by a distributed lag of past realized volatilities.[17] The estimation of an infinite number of b_i coefficients implicit in this representation is, of course, not practically feasible, and the different risk models in effect represent alternative ways of restricting the $b(L)$ lag polynomial to allow for its meaningful estimation, as exemplified by the RV-based ARIMA model originally proposed by Andersen, Bollerslev, Diebold, and Labys (2003a), and the MIDAS model advocated by Ghysels, Santa-Clara, and Valkanov (2006) in which $b(L)$ is parameterized in terms of Beta functions.

The notion of multiple volatility components, or factors, is also commonly used for parsimoniously representing the $b(L)$ lag polynomial. The HAR model of Corsi (2009), for

ularly scheduled U.S. macroeconomic news announcement as non-farm payroll, durable goods orders, retail sales, housing starts, and the Philadelphia FED's business outlook survey.

[17] In addition to the risk models nested in this AR(∞) representation, we also explored models that decompose the daily RV's into continuous and jump components (Andersen, Bollerslev, and Diebold, 2007), and models that differentiate between up and down realized semi-variances (Patton and Sheppard, 2015). For the monthly forecast horizon primarily analyzed below, none of these alternative RV measures and models resulted in superior out-of-sample forecasts. Chen and Ghysels (2011) also report largely symmetric "news impact curves" for U.S. aggregate equity indexes at the one-month horizon.

example, is based on a weighted sum of a daily, weekly and monthly volatility factors.[18] In this situation, what ultimately matters from a practical forecasting perspective is the factors' ability to capture the influence of the lagged RV's. Ideally, we want a set of factors that "span" the $b(L)$ lag space well, while still enforcing some commonality and "smoothness," thereby enabling a unified set of factors to be used for all assets and asset classes by simply altering the weights of the different factors.[19]

Regardless of the way in which the $b(L)$ lag polynomial is parameterized, the b_i coefficients are usually estimated on an asset-by-asset basis. This ignores the cross-asset similarities in the dynamic dependencies discussed in the previous section. In addition to the standard individual asset-by-asset OLS-based estimation of the models, we therefore also explore the use of panel regression techniques that force the coefficients to be the same within and across different asset classes. As demonstrated below, doing so imbues the resulting risk model with a built-in robustness and statistically significant superior out-of-sample forecast performance.

4.2. "Centering": Eliminating the Level Parameter in a Robust Fashion

Even though one might naturally restrict the dynamic $b(L)$ lag coefficients to be the same across assets to exploit the commonalities in the dynamic dependencies and distribution of the standardized volatilities, $RV_t/Mean(RV_t)$, the very different volatility levels for different asset classes means that it is unreasonable to force the b_0 intercepts to be the same. To circumvent this, and allow for meaningful cross-asset estimation, we "center" the risk models by replacing the intercept with a long-run volatility factor RV_t^{LR}, equal to the expanding sample mean of the daily RV's from the start of the sample up until day t. Forcing all of the b_i coefficients to sum to one, including the coefficient for the RV_t^{LR} factor, ensures that the iterated long-run forecasts from the model constructed on day t converges to this day t estimate of the "unconditional" volatility.[20]

Although seemingly complicated to implement, this "centering" of the risk models is easily enforced by subtracting the long-run volatility factor from all of the RVs, including

[18] Related representations based on the combination of multiple distributed lag polynomials have also been proposed in the MIDAS literature and elsewhere; see, e.g. Ghysels, Sinko, and Valkanov (2007).

[19] For instance, the data might suggest a lag structure with a 10-day Center-of-Mass (CoM) for natural gas futures, but a 100-day CoM for Japanese 10-year bond futures. By assigning different weights to the same set of unified factors with different CoMs, the estimated coefficients can effectively tilt towards a risk model on the faster end for natural gas and towards a model on the slower end for Japanese bond futures.

[20] This mirrors the idea of variance targeting in GARCH models proposed by Engle and Mezrich (1996), in which the intercept in the conditional variance equation is replaced by a (scaled) estimate of the long-run "variance target" to which the forecasts converge. It also resembles the Spline-GARCH model of Engle and Rangel (2008), and the use of a low-frequency volatility component to scale the forecasts.

the right-hand-side RV forecast target,

$$RV_{t+1} - RV_t^{LR} \;=\; b_1(RV_t - RV_t^{LR}) \;+\; b_2(RV_{t-1} - RV_t^{LR}) \;+\; \ldots \;+\; \epsilon_t. \tag{3}$$

When the regression is run in this way, the coefficients are free (i.e., need not sum to one), but, if we collect terms for RV_t^{LR} on the right-hand-side, then RV_t^{LR} has an implied coefficient of $1 - b_1 - b_2 - \ldots = 1 - b(1)$ such that all the implied coefficient do sum to one. By eliminating the level of the volatility, this alternate representation allows for the meaningful estimation of common dynamic b_i coefficients by panel regression techniques. More complicated Bayesian shrinkage type procedures, in which the estimated coefficients are allowed to differ across assets could, of course, be applied. However, we purposely restrict the coefficients to be the same within asset classes or across all assets, to allow for a direct comparison with the individually estimated risk models.[21]

4.3. Multi-Period and Other Volatility Forecasts

The AR(∞) model in (2) and the centered version thereof in (3) are directly geared to forecasting the one-day-ahead variance. Longer-run forecasts, say over weekly or monthly horizons, may be obtained by recursively substituting the forecasts for the future daily RV's into the right-hand-side of the model, subsequently adding up the resulting one, two, three, etc. days-ahead forecasts to achieve the multi-period RV forecast over the requisite horizon. Instead, a much simpler approach for constructing multi-day-ahead forecasts is to estimate the risk model as such from the start. That is, by replacing the daily variance RV_{t+1} on the left-hand-side of the risk model, with the realized variance over the forecast horizon h of interest, say $RV_t^h \equiv \frac{1}{h}\sum_{i=1}^{h} RV_{t-h+i}$. In the forecasting literature, this approach is commonly referred to as direct as opposed to iterated forecasts.[22]

In particular, for the "monthly," or 20-day, forecast that we focus on below, and the generic risk model in (2), we have

$$RV_{t+h}^h \;=\; b_0^h \;+\; b^h(L)RV_t \;+\; \epsilon_t^h, \tag{4}$$

with $h = 20$. The b_i^h coefficients, which dictate the "speed" of the model, will obviously depend on the forecast horizon, as indicated by the superscripts h. For notational simplicity,

[21]We also experimented with the use of common long-run asset-class RV_t^{LR}s, but we did not find that these models resulted in obviously superior forecasts.

[22]A similar distinction has been made in the context of option pricing and the use of exactly matched-by-horizon versus cumulated volatility estimates; see e.g., Karolyi (1993). This issue has also been extensively studied in the MIDAS literature; see, e.g., Ghysels, Rubia, and Valkanov (2009).

however, we will drop this superscript in the sequel. Also, even though ϵ_t^h will generally be serially correlated up to order $h-1$, we will simply denote the residuals in all of the models discussed below as ϵ_t for short.[23] In theory, if the model for the one-day-ahead RV_{t+1} in (2) is correctly specified, the iterated forecasts from that model would be the most efficient. However, there is ample empirical evidence that even minor model mis-specifications often get amplified in iterated volatility forecasts, and as a result the direct forecasts produced from a model such as (4) are typically superior in practice; see, e.g., Andersen, Bollerslev, Diebold, and Labys (2003a), Ghysels, Rubia, and Valkanov (2009), and Sizova (2011).[24]

This same basic idea may also be used for forecasting other functions of the variance, by simply replacing RV_{t+h}^h on the left-hand-side in equation (4) with the volatility object of interest. For instance, the future volatility as opposed to the variance, or the inverse of the variance, are often of primary import. Unless the volatility is constant, or perfectly predictable, simply transforming the forecast for the variance will result in a systematically biased forecast.[25] We will briefly revisit this issue in our discussion of the empirical results below. However, we begin with a brief discussion of the specific risk models, old and new, that we rely on.

4.4. HAR Models

The original HAR model of Corsi (2009) has proven very successful. It may be succinctly expressed as,

$$RV_{t+h}^h = \beta_0 + \beta_D RV_t + \beta_W RV_t^W + \beta_M RV_t^M + \epsilon_t, \tag{5}$$

where RV_t^W and RV_t^M denote the 5-days (weekly) and 20-days (monthly) realized volatilities, respectively, thus implying a step function for the b_i coefficients in the omnibus AR(∞) representation. Corsi (2009) finds that from a forecasting perspective the approximate long-memory HAR model performs on par with or better than the ARIMA model used by Andersen, Bollerslev, Diebold, and Labys (2003a), and the HAR model has now also emerged as somewhat of a benchmark in the financial econometrics literature for judging other RV-based forecasting procedures.[26] Thus, to allow for a direct comparison with the existing literature,

[23]The use of overlapping daily data in the estimation of the models also means that the conventional standard errors for the coefficient estimates will have to be adjusted to account for the overlap; see, e.g., Hansen and Hodrick (1980) and Newey and West (1987).

[24]Note that Ghysels, Rubia, and Valkanov (2009) refer to the direct forecasts as MIDAS.

[25]By a standard first-order Taylor series expansion, $E(\sqrt{X}) \cong E(X)^{1/2} - \frac{1}{8}Var(X)$. Hence, $[E_t(RV_{t+h}^h)]^{1/2}$ is an upward biased forecast of the future volatility. Similarly, $E(1/X) \cong 1/E(X) + Var(X)/E(X)^3$, so that one over $E_t(RV_{t+h}^h)$ will result in a downward biased forecast for the relevant ratio.

[26]Analogous multi-factor formulations have previously been used as way to approximate long-memory dynamic dependencies in the context of parametric stochastic volatility models; see, e.g., Gallant, Hsu, and

we rely on the standard "uncentered" version of the HAR model. However, in addition to the results from individual asset-by-asset estimation, we also report the results from fixed effect panel-based estimation, in which we restrict the β_D, β_W and β_M coefficients to be the same, but allow the β_0 coefficients to differ across different assets. Our forecast comparison tests for the differently estimated models provide an indirect assessment of these restrictions.

The daily, weekly and monthly volatility factors in the original HAR model were initially justified by the idea that volatility changes in response to economic events that occur, and agents that operate, on these specific time-scales. To allow for a more flexible lag structure, we extend the model to include the first six daily lagged RV's with their own individually estimated coefficients.[27] Additionally, to allow for a direct impact of changes in the volatility beyond the monthly horizon, we also include an annual volatility factor RV_t^A, defined by the (normalized) summation of the daily RV's over the past 261-days. Further "centering" the model results in the formulation,

$$RV_{t+h}^h - RV_t^{LR} = \sum_{j=1}^{6} \beta_j (RV_{t+1-j} - RV_t^{LR}) + \beta_M (RV_t^M - RV_t^{LR}) + \beta_A (RV_t^A - RV_t^{LR}) + \epsilon_t. \quad (6)$$

We will refer to this as the (centered) HAR-Free model in the sequel.

The step-wise nature of the volatility factors employed in the HAR models, imply that the forecasts from the models are subject to potentially abrupt changes as an unusually large/small daily lagged RV drops out of the sums for the longer-horizon lagged volatility factors. The added flexibility afforded by the HAR-Free model further renders the forecasts constructed from that model potentially more erratic and susceptible to "noise," especially when estimated on an individual asset-by-asset basis over a relatively short sample. Our remaining risk models rely on alternative $b(L)$ polynomials for "smoothing" out these problems.

4.5. MIDAS Models

The original HAR model may be interpreted as a special case of MIDAS regressions with step functions (see, e.g., the discussion in Andersen, Bollerslev, and Diebold, 2007; Ghysels, Sinko, and Valkanov, 2007; Corsi, 2009), while the HAR-Free model is closely related to the

Tauchen (1999). The use of monthly lagged volatility in the daily ARCH model estimated by French, Schwert, and Stambaugh (1987) also provides an early precedent to this formulation.

[27]The empirical evidence in Craioveanu and Hillebrand (2012) pertaining to a set of individual equities, suggests that the use of a more flexible lag structure does not result in improved out-of-sample forecasts compared to the original HAR model in (5). However, the lasso-based model selection procedures in Audrino and Knaus (2016) generally do not recover the exact lag structure of the original HAR model, although the out-of-sample performance of the selected models appear indistinguishable from that of the original HAR.

so-called U-MIDAS model (Foroni, Marcellino, and Schumacher, 2015). In contrast to HAR models, however, one of the main objectives of the MIDAS approach is the specification of "smooth" distributed lag polynomials for representing the dynamic dependencies. Another main theme of the MIDAS literature, of course, relates to the use of data sampled at different frequencies, and the choice of sampling frequency for the regressor variables. Addressing both of those issues, Ghysels, Santa-Clara, and Valkanov (2006) conclude that the direct modeling of high-frequency data does not result in systematically better volatility forecasts compared to the forecasts from a model of the form in (2) based on the daily RVs only.[28] They also propose a specific parameterization for the $b(L)$ lag polynomial based on beta functions. This representation has now emerged as somewhat of a standard in the MIDAS literature.

Thus, directly following Ghysels, Santa-Clara, and Valkanov (2006), we implement a MIDAS model of the form,

$$RV_{t+h}^{h} \;=\; \beta_0 \;+\; \beta_1[a(1)^{-1}a(L)]RV_t + \epsilon_t, \tag{7}$$

in which the non-zero coefficients in the $a(L)$ lag polynomial are defined by scaled beta functions,

$$a_i \;=\; \left(\frac{i}{k}\right)^{\theta_1-1}\left(1-\frac{i}{k}\right)^{\theta_2-1}\Gamma(\theta_1+\theta_2)\Gamma(\theta_1)^{-1}\Gamma(\theta_2)^{-1}, \qquad i=1,...,k, \tag{8}$$

where $\Gamma(\cdot)$ denotes the Gamma function. The normalization by $a(1) \equiv a_1+...+a_k$ in equation (7) ensures that the coefficients in the $[a(1)^{-1}a(L)]$ lag polynomial sum to unity, so that the β_1 coefficient is uniquely identified. Implementation of the model still requires a choice of the cutoff k, and the two tuning parameters θ_1 and θ_2. Again, directly mirroring Ghysels, Santa-Clara, and Valkanov (2006) we fix the cutoff at $k = 50$, and set the tuning parameter $\theta_1 = 1$. This choice of θ_1 is now also commonly employed in the MIDAS literature more generally. Lastly, following the approach of Ghysels and Qian (2016), we determine the remaining θ_2 tuning parameter by a grid search, in which we profile the predictive R^2 from the model as a function of θ_2 together with the freely estimated β_0 and β_1 parameters, choosing the value of θ_2 that maximizes the predictability over the full sample.[29]

[28]Echoing this conclusion, Clements, Galvao, and Kim (2008) also find that that the parameterization implicit in the original HAR model generally results in superior out-of-sample forecasts compared to a freely parameterized MIDAS model for the high-frequency intraday squared returns.

[29]In addition to this benchmark MIDAS model, we also experimented with a series of more elaborate specifications based on the mixture of multiple beta polynomials. Further details concerning these additional models are available in the Supplementary Appendix.

4.6. HExp Models

The "smooth" beta polynomial employed in the MIDAS model avoids the step-wise changes inherent in the forecast from the HAR component-type structure. Further extending this idea, our last set of risk models rely on a mixture of "smooth" Exponentially Weighted Moving Averages (EWMA) of the past realized volatilities. Simple EWMA filters with a pre-specified center of mass are often used in practice. Instead, we explicitly estimate the relative importance of different EWMA factors constructed from the past daily RV's,

$$ExpRV_t^{CoM(\lambda)} \equiv \sum_{i=1}^{500} \frac{e^{-i\lambda}}{e^{-\lambda} + e^{-2\lambda} + ... + e^{-500\lambda}} RV_{t+1-i}, \tag{9}$$

where λ defines the decay rate of the weights, and $CoM(\lambda)$ denotes the corresponding center-of-mass, $CoM(\lambda) = e^{-\lambda}/(1 - e^{-\lambda})$.[30] The center-of-mass of each exponential RV measure effectively summarizes the "average" horizon of the lagged realized volatilities that it uses. Conversely, for each center-of-mass, we can compute the corresponding rate of decay as $\lambda = \log(1 + 1/CoM)$. Hence, we can think of $\lambda = \log(1 + 1/125) = 0.008$ as an annual $ExpRV$ risk measure because the corresponding center of mass is 125 trading days, that is, about half a year, just like an annual equal-weighted average realized volatility, RV_t^A.

Correspondingly, we focus on similar horizons to the ones used in the HAR-Free model. However, we rely on the exponential RV's to "span" this universe of past RV's in a way that is both parsimoneous and "smooth," mixing four $ExpRV_t^{CoM(\lambda)}$ factors with λ chosen to equate the center-of-mass to 1, 5, 25 and 125 days, respectively. Further "centering" the model around the expanding long-run volatility factor, we obtain the following new risk model,

$$RV_{t+h}^h - RV_t^{LR} \;=\; \sum_{j=1,5,25,125} \beta_j (ExpRV_t^j - RV_t^{LR}) \;+\; \epsilon_t. \tag{10}$$

We will refer to this specification as the Heterogeneous Exponential, or HExp, model. This model, of course, is still nested in the omnibus distributed-lag model in (3). As such, it may also be interpreted as another MIDAS specification. Importantly, however, the use of pre-specified volatility factors for characterizing the volatility dynamics that do not depend on any unknown parameters, implies that the model is straightforward to estimate by standard OLS

[30] The center of mass is formally defined as the weighted-average time period for the lags used,

$$CoM(\lambda) \equiv \frac{\sum_{t=0}^{\infty} e^{-\lambda t} t}{\sum_{t=0}^{\infty} e^{-\lambda t}} = \frac{e^{-\lambda}}{1 - e^{-\lambda}},$$

where we ignore that the sum in (9) only uses the first 500 lags to achieve a simple formula (the influence of the remaining lags is numerically immaterial).

on an asset-by-asset basis, or panel regression procedures that restrict the beta coefficients to be the same across groups of assets.[31]

Motivated by the cross-asset and cross-market volatility spillover effects discuss in Section 3, our final risk model augments the asset-specific HExp model in (10) with the global risk factor $ExpGlRV_t^5$. Specifically, $ExpGlRV_t^5$ is the 5-day center-of-mass EWMA of the realizations of the global risk factor $GlRV_t$ defined in Section 3.2, yielding the risk model:

$$RV_{t+h}^h - RV_t^{LR} \;=\; \sum_{j=1,5,25,125} \beta_j (ExpRV_t^j - RV_t^{LR}) \;+\; \beta_5^{Gl}(ExpGlRV_t^5 - RV_t^{LR}) \;+\; \epsilon_t, \quad (11)$$

We note that, as discussed in Section 3.2, the global risk factor is defined on an asset specific basis to avoid any overlap between RV_t and $GlRV_{t-1}$, and further normalized to have the same time t expanding sample mean as the specific asset. The inclusion of the global risk factor naturally enforces a degree of commonality over-and-above that afforded by restricting the beta coefficients to be the same. We will refer to model (11) as the HExpGl.

5. Risk Inference: Model Estimation and Forecasting

We will focus our discussion on a one-month forecast horizon. We report both in-sample results, in which we rely on all the available data, as well as out-of-sample forecasts, in which we estimate the models based on an expanding window of the data up to that point in time.[32] We rely on the explained sum of squares divided by the total sum of squares within and across asset classes as way to succinctly summarize the performance of the different models. To allow for meaningful comparisons between the in- and out-of-sample results, we use the expanding long-run sample mean, or the RV_t^{LR} factor, in the calculations of the out-of-sample R^2's.[33] All of the models are estimated by OLS on an individual asset-by-asset basis, by panel regressions that restrict the coefficients to be the same for all of the assets

[31] We also experimented with related specifications based on linearly and hyperbolically decaying volatility factors, mixtures of multiple beta polynomials as the one employed in the MIDAS model, as well as mixtures of both exponentially and hyperbolically decaying factors. For the monthly forecast horizon primarily analyzed below, none of these alternative models resulted in systematically superior forecasts compared to the HExp model in (10). Further details concerning some of these additional results are available in the Supplementary Appendix.

[32] We require that an asset has at least one full calendar year of RV's before it is included in the estimation.

[33] The out-of-sample R^2 is formally defined as $R^2 = 1 - \sum_{t=1}^{T}(RV_{t+20}^M - \widehat{RV_{t+20}^M})^2 / \sum_{t=1}^{T}(RV_{t+20}^M - RV_t^{LR})^2$, where $\widehat{RV_{t+20}^M}$ refer to the predictions from one of the risk models. This mirrors the out-of-sample R^2s commonly used in evaluating return predictability; see, e.g., Campbell and Thompson (2008). The HAR and MIDAS models in a few instances produce very large out-of-sample forecasts. To avoid deflating the corresponding R^2's for these models and allow for empirically more meaningful comparisons, we follow Swanson and White (1997) in applying an "insanity filter," in which we replace any forecast that exceeds the maximum RV_t^M observed up to that point with RV_t^{LR}; i.e., "insanity" is replaced by "ignorance."

within an asset class, and by panel regressions that restrict the coefficients to be the same across *all* assets.[34] We refer to these alternative estimation schemes as "Individual Asset," "Panel," and "Mega" panel, respectively.

5.1. Basic Estimation and In-Sample Forecasting Results

We begin our discussion by considering the in-sample estimation results. The different specifications of the models complicate any direct comparisons of the estimated β coefficients. However, all of the risk models, except for the HExpGl model, are nested in the univariate AR(∞) representation in (2). Hence, whereas the estimated β coefficients are not directly comparable, the dynamics of the different risk models may be meaningfully compared in terms of the implied b_i coefficients in that representation. To this end, Figure 6 depicts the impled $b(L)$ polynomials out to a lag length of 25-days for the mega-panel-based estimated models.[35] For comparison purposes, we also include the weights for a "monthly" equally weighted 21-Day RV corresponding to a simple random walk type forecast, or equivalently a HAR model with $\beta_M = 1$ and $\beta_0 = \beta_D = \beta_W = 0$. As the figure shows, with the exception of the flat weights for the 21-Day RV, the estimates are generally fairly close. Nonetheless, the HAR-Free and HExp models both appear slightly "faster" than the MIDAS model, with a more rapid initial decay and less weight assigned to the intermediate lags ranging between five days and two weeks.

The figure also visualizes the step-wise nature of the implied b_i coefficients for both of the HAR models. As a result, forecasts constructed from these models are more susceptible to abrupt changes, and therefore potentially also more costly to implement, than the forecasts from the "smoother" risk models. For instance, looking at the daily change in the in-sample forecasts averaged across all assets, the sample variance $Var(RV^M_{t+21|t+1} - RV^M_{t+20|t})$ equals $2.11 \cdot 10^{-4}$ for the original HAR model, compared to $1.37 \cdot 10^{-4}$ for the HExp model. Moreover, the first-order autocorrelation of the change in the daily forecasts equals -0.21 for the HAR model, compared to -0.01 for the HExp model, suggesting that the latter risk model would be both easier and cheaper to implement in practice. We will return to this issue and the "speed" of the models in our discussion of the utility-based comparisons in Section 6 below.

[34] As noted above, the additional θ_2 parameter needed for the MIDAS model is estimated by profiling. Since the R^2s are fairly "flat" over a wide range of θ_2s, we rely on the full in-sample estimates for the out-of-sample analysis as well, implicitly ignoring any look-ahead biases. For the single MIDAS model that restrict the parameters be the same across *all* assets this results in a common $\hat{\theta}_2 = 6.50$. Further details concerning the θ_2 MIDAS estimates are available in the Supplementary Appendix.

[35] Similar graphs for the individually estimated risk models for the four representative assets shown in Figures 1 and 2, and the panel-based estimation of the four asset classes, are available in the Supplementary Appendix.

Meanwhile, the in-sample predictive R^2's for the different risk models, together with the 21-Day RV benchmark, are given in Table 4. The individually estimated models by construction always result in larger R^2's than the panel-based estimation of the asset class specific risk models, and the mega risk models that restrict the coefficients to be the same across all assets. Interestingly, however, with the exception of the HExpGl model, the differences in the in-sample R^2's tend to be fairly small across the different risk models, especially for the individually estimated models. The R^2's for all of the risk models, whether individually or panel-based estimated, also easily exceed the R^2's for the 21-Day RV random walk forecasts.

The Diebold and Mariano (1995) (DM) tests for comparing predictive accuracy, reported in the bottom panel of the table, further corroborate these observations. Specifically, taking the individually estimated HExp model as the benchmark, we formally test the null hypotheses of equal predictive ability by calculating robust t-statistics for the sample means of the time series comprised of the average standardized (by the mean of the realized variation) squared error losses for each of the different models, estimation procedures, and asset classes.[36] With the exception of the predictions pertaining to the foreign exchange market and the predictions from the HExpGl model, the majority of the t-statistics are significant at the usual 5% level. Looking across the different estimation procedures, the by-construction higher in-sample R^2s for the individually estimated models also generally translate into statistically significant lower losses compared to the panel and mega models that restrict the coefficients to be the same across groups of assets. As a case in point, the t-statistic for the individually estimated HExp model versus the mega HExp model for all assets equals -2.21. Among all of the mega-based models, only the mega HExpGl model does not results in statistically inferior in-sample predictions compared to the individually estimated HExp model.

5.2. Out-of-Sample Forecasting Results

This systematic ordering of the individual versus panel-based estimated models based on the in-sample results in Table 4, is essentially reversed for the out-of-sample results reported in Table 5.[37] The mega-panel estimation that restricts the coefficients to be the same across

[36]The standardization of the squared error loss allows for a more meaningful comparison of the losses across the different assets. It also reflects the empirical observation that the normalized distributions of the realized volatilities are similar across the different assets, and corresponds more directly to the scale-invariant R^2s reported in the top panel. However, were similar DM-test statistics are obtained for the average non-standardized loss differentials; further details are reported in the Supplementary Appendix.

[37]To conserve space we only report the out-of-sample results for all of the assets combined. The specific results for the different asset classes are included in the Supplementary Appendix. The slightly higher out-of-sample R^2's compared to the in-sample results stems from the different demeaning factors used in the denominators of the R^2s.

all assets now typically results in the highest R^2's. Looking across the different columns, the out-of-sample results also reveal a much clearer rank ordering among the different risk models, with the HExp and HExpGl models performing the best overall.

The DM-tests based the average out-of-sample standardized squared error losses, reported in the bottom part of the table, again corroborate these conclusions.[38] Now taking the mega HExp model as the benchmark, the pairwise tests show that the forecasts from that model result in significantly lower losses than the forecasts from *all* of the other models, except for the different HExpGl models. Importantly, the forecasts from the mega HExp model also result in significantly lower losses than the forecasts from the individually estimated HExp models and panel HExp models that allow for different coefficients across asset classes. Intuitively, the mega estimation approach provides a built-in robustness against influential outliers, and in turn superior out-of-sample forecast.

To further appreciate this point, the expected squared forecast error loss may be expressed as the sum of the squared forecast bias, the variance of the forecasts, plus the variance of the "irreducible error" associated with the true (unknown) conditional expectation $RV_{t+h}^h - E_t(RV_{t+h}^h)$; see, e.g., the discussion in Hastie, Tibshirani, and Friedman (2009). Looking at the bias-variance trade-off implicit in the squared forecast error losses thus help explain why the pooling and panel-based estimation that exploit the commonalities and reduce the parameter estimation error uncertainty generally works the best from an out-of-sample forecasting perspective and result in the highest predictive R^2s. In particular, while the squared out-of-sample forecast biases are very small for all of the different models and estimation methods and effectively immaterial, the individually estimated risk models systematically result in the most variable forecasts by quite a wide margin. For the HExp model, for example, the average variance of the forecasts is reduced by almost 30% for the mega model that restrict the coefficients to be the same for all assets compared to the average forecast variance for the individually estimated HExp models; more detailed results along these lines for all of the different models are provided in the Supplementary Appendix. Importantly, by the same reasoning, this does not necessarily imply that the best performing forecasting model is somehow closest to the "true" model, only that more parsimonious risk models tend to produce better out-of-sample forecasts.

[38] Note that the out-of-sample DM-tests are not formally justified as tests for the correct model specification, as the magnitude of the parameter estimation errors decrease with the size of the expanding estimation window, thereby rendering the losses non-stationary. Nonetheless, taking the model forecasts as the primitives, the tests may still by justified as tests for how competing risk forecasters using the different models would have fared; see also the discussion in Diebold (2015).

5.3. *Robustness and Additional Empirical Results*

The main empirical findings discussed above were based on "centered" risk models, forecasting the realized variation one month ahead, using the lagged RVs for the whole day. We now briefly discuss the sensitivity to each of these assumptions and modeling choices. More detailed results and tables related to these additional empirical investigations and robustness checks are provided in the Supplementary Appendix.

5.3.1. *"Centered" versus "Uncentered" Risk Model Forecasts*

Our "centering" of the HAR-Free and HExp models formally alter the models relative to their "uncentered" counterparts. Specifically, rewriting the generic "centered" risk model in (3) with only RV_{t+1} on the left-hand-side, the "intercept" on the right-hand-side becomes $(1 - b_1 - b_2 - \ldots)RV_t^{LR} = (1 - b(1))RV_t^{LR}$. This time-varying "intercept" obviously differs from the time-invariant b_0 intercept in the more standard risk model (2). Only by using the unconditional sample variance for the full sample, say RV_T^{LR}, in place of the expanding RV_t^{LR}, would the full in-sample OLS and panel-regression estimates of the dynamic b_i coefficients be the same for the two models.

Meanwhile, the information structure implicit in the "centered" risk model is more in-tune with the practical uses of the model, in that only information as of day t is used for the construction of the day $t+1$ risk forecast. Importantly, this translates into the actual forecast performance of the differently formulated models. For instance, while the "uncentered" versions of the HAR-Free and HExp models both work better than the "centered" versions in-sample with 0.8% higher R^2s for the mega-panel-based estimation, the out-of-sample R^2s for the same mega-based models are 0.6% and 0.7% higher, respectively, for the "centered" models. This same general conclusion holds true for all of the other models.

5.3.2. *Other Forecast Horizons*

The results discussed so far all pertain to monthly predictions. We also implemented the same risk models for both shorter (daily and weekly) and longer (60-days) prediction horizons. Consistent with the results for the monthly predictions, the differences in the in-sample R^2s are generally fairly small across the different risk models. Also, while the individually estimated models necessarily result in the highest in-sample R^2's, the mega-based estimation that restricts the coefficients to be the same across all assets typically results in the best performing models out-of-sample. Interestingly, the largest (in a relative sense) improvements in the R^2's compared to the original HAR model manifest over longer horizons, indirectly underscoring the importance of including longer-run volatility factors and properly "centering" the risk models when forecasting further into the future.

5.3.3. Risk Models based on Daily Returns vs. Intraday Realized Volatilities

As noted above, the monthly out-of-sample predictive R^2's for the RV-based risk models reported in Table 5 easily exceed those based on monthly past realized volatilities constructed from daily squared returns, or simple EWMA of the lagged daily squared returns. To further investigate whether these improvements arise from better risk modeling or better risk measurements, or both, we implement the identical suite of risk models and predictive R^2s using the daily squared returns in place of the daily RVs. Not surprisingly, benchmarking the performance of the models against the sum of the daily squared returns results in more "noisy" and systematically lower R^2s. More important, the risk models based on realized volatilities systematically outperform the identical risk models based on risk factors constructed from daily squared returns. The out-of-sample R^2 for the mega HExp model, for example, improves from 44.0% with the daily squared returns to 47.3% with daily RVs, directly underscoring the informational advantages in risk modeling afforded by the use of the more accurate daily realized volatility measures.

5.3.4. Intraday and Overnight Risk Models

As discussed in Section 2.3, there is ample empirical evidence dating back at least to French and Roll (1986) that volatility tend to be lower during non-trading hours than during the active part of the trading day.[39] The daily RVs employed in our risk models simply add the intraday variation to the overnight squared returns to obtain an estimate of the volatility for the whole day. However, different dynamic dependencies may be at work intraday and overnight. To investigate whether implicitly restricting the dynamics to be the same result in inferior forecasts, we implement separate risk models for the intraday and overnight components. Our results confirm that it is generally more difficult to predict the overnight volatility compared to the volatility for the active part of the trading day. Meanwhile, the HExp model that performed the best for predicting the daily RVs, also performs the best for separately predicting the intraday and overnight volatility components. Importantly, combining the forecasts from the separately estimated models for the intraday and overnight volatilities to

[39]French and Roll (1986) argue that private information revealed through the process of trading may help explain these differences. Recent work by Boudoukh, Feldman, Kogan, and Richardson (2016) using textual analysis to identify specific types of public "news," find that the ratio of overnight variation to within day variation also varies importantly with the intensity of the "news," suggesting that public information account for a non-trivial portion of the return variation as well.

obtain a forecast for the whole day, or month, do not result in obviously superior out-of-sample forecasts compared to the risk models for the daily RVs. For example, combining individually estimated HExp models for the intraday and overnight volatilities results in an average out-of-sample R^2 of 46.8%, compared to the average R^2 of 47.3% for the single mega HExp model.

5.3.5. Volatility vs. Variance Forecasting

Investment decisions and risk measurements often depend on functions of the future variance, as opposed to the future variance itself. As discussed in Section 4.3, the different risk models are easily modified to accommodate this by replacing RV_{t+h}^{h} on the left-hand-side with the volatility object of interest. In so doing, the RVs and variance risk factors on the right-hand-side in the models are also naturally replaced by the same transformations. Implementing, such transformed versions of the risk models for forecasting $\sqrt{RV_{t+h}^{h}}$ again confirms our general conclusions. While the individually estimated risk models necessarily result in the highest R^2's in-sample, the mega models that restrict the dynamic coefficients to be the same systematically result in the best out-of-sample volatility forecasts.

The practical uses of risk models, of course, face a host of other issues and tradeoffs related to the actual costs and benefits of implementing the forecasts from the models. To illustrate these issues, we turn next to a utility-based framework for evaluating the benefits of an investment strategy involving the notion of equal risk shares. To keep the models simple and the results directly comparable to the ones discussed in the previous sections, we purposely focus on the risk models estimated to forecast the variance.

6. Risk Models in Action: Quantifying the Utility Benefits

This section further highlights the benefits of dynamic risk modeling, by considering a simple utility-based framework: an investor with mean-variance preferences investing in an asset with time-varying volatility and a constant Sharpe ratio. In contrast to the approach of Fleming, Kirby, and Ostdiek (2001, 2003) subsequently used by many other studies, which depends on forecasts for both returns and volatilities, our framework only depends on volatility forecasts.

6.1. Expected Utility and Risk Targeting

By standard arguments, the expected utility may, up to a factor of proportionality, conveniently be approximated as (dropping constant terms that only depend on time-t variables),

$$E_t(u(W_{t+1})) = E_t(W_{t+1}) - \frac{1}{2}\gamma^A Var_t(W_{t+1}), \tag{12}$$

where $\gamma^A \equiv -u''/u'$ denotes the absolute risk aversion of the investor. Assume that the investor allocates a fraction x_t of his current wealth to a risky asset with return r_{t+1} and the rest to a risk-free money market account earning r_t^f. His future wealth becomes $W_{t+1} = W_t(1 + x_t r_{t+1} + (1 - x_t)r_t^f) = W_t(1 + r_t^f) + W_t x_t r_{t+1}^e$, where $r_{t+1}^e \equiv r_{t+1} - r_t^f$ denotes the excess return, resulting in an expected utility of (again dropping constant terms),

$$\begin{aligned} U(x_t) &= W_t\left(x_t E_t(r_{t+1}^e) - \frac{\gamma}{2}x_t^2 Var_t(r_{t+1}^e)\right) \\ &= W_t\left(x_t E_t(r_{t+1}^e) - \frac{\gamma}{2}x_t^2 E_t(RV_{t+1})\right), \end{aligned} \tag{13}$$

where $\gamma \equiv \gamma^A W_t$ refers to the investor's relative risk aversion.

To focus on risk modeling, we assume that the conditional Sharpe ratio, defined as $SR \equiv E_t(r_{t+1}^e)/\sqrt{E_t(RV_{t+1})}$, is constant.[40] Under this assumption, the expected utility simply depends on the position x_t, together with the expected realized volatility $E_t(RV_{t+1})$,

$$U(x_t) = W_t\left(x_t SR\sqrt{E_t(RV_{t+1})} - \frac{\gamma}{2}x_t^2 E_t(RV_{t+1})\right). \tag{14}$$

The optimal portfolio that maximizes this utility is obtained by investing the fraction of wealth $x_t^* = E_t(r_{t+1}^e)/(\gamma E_t(RV_{t+1}))$ in the risky asset, or alternatively,

$$x_t^* = \frac{SR/\gamma}{\sqrt{E_t(RV_{t+1})}}, \tag{15}$$

resulting in an expected utility of

$$U(x_t^*) = \frac{SR^2}{2\gamma}W_t. \tag{16}$$

In other words, the investor optimally targets a volatility of SR/γ, since the conditional

[40] This assumption naturally corresponds to our modeling of the realized variation in the previous sections. Alternatively, one might assume that $E_t(r_{t+1}^e/\sqrt{RV_{t+1}})$ is constant, resulting in a Jensen's inequality adjustment term $E_t(\sqrt{RV_{t+1}})/E_t(RV_{t+1})$ for the optimal investment positions and expected utilities derived below. This same general setup also generalize to time-varying conditional Sharpe ratios, $SR_t \equiv E_t(r_{t+1}^e)/\sqrt{E_t(RV_{t+1})}$, as long as SR_t is independent of the risk $E_t(RV_{t+1})$.

standard deviation of the x_t^* portfolio equals $\sqrt{Var_t(x_t^* r_{t+1}^e)} = SR/\gamma$. When the predicted volatility $\sqrt{E_t(RV_{t+1})}$ is above the "risk target" SR/γ, the agent only invests part of his wealth in the risky asset ($x_t^* < 1$). Conversely, when the predicted risk is below the target, the investor applies leverage ($x_t^* > 1$) to reach his risk target. This behavior mimics in a simple way the actual trading behavior of many hedge funds with explicit volatility targets and so-called risk parity investors.

For concreteness, and in parallel to the forecasting results discussed in the previous section, we focus on a monthly forecast horizon. Guided by the typical results reported in the extant investment literature, we take the corresponding Sharpe ratio and coefficient of risk aversion to be $SR = 0.3$ and $\gamma = 3$, respectively (see, e.g., Pedersen, 2015, for empirical evidence pertaining to the same broad set of assets analyzed here).[41] These assumptions in turn imply that the investor's optimal position equals,

$$x_t^* = \frac{10\%}{\sqrt{E_t(RV_{t+1})}}, \tag{17}$$

or equivalently that the investor optimally targets a volatility of 10%. By direct substitution in (16) above, the associated utility for this optimally targeted portfolio equals,

$$U(x_t^*) = 1.5\%\, W_t, \tag{18}$$

meaning that the investor would be willing to give up 1.5% of his wealth to have access to the x_t^* portfolio rather than simply investing in the risk-free asset. Put differently, since the utility from (14) of a risk-free position equals $U(0) = 0$, the investor would receive the same utility by either: (i) trading the risky asset optimally while paying a fee of 1.5% times his wealth; or (ii) putting all of his money in the risk-free asset.

To further appreciate this number, consider the expected return of the investor's strategy. Given a Sharpe ratio of 0.3 and a 10% risk target, the investor expects to make an excess return of 3%. However, at the optimally targeted position, half of this return is "lost" due to the dis-utility of risk, so the investor is left with a benefit of only 1.5%.[42]

To explicitly quantify the utility gains from different risk models, let $E_t^\theta(\cdot)$ denote the expectations from model θ. Also, let $E_t(\cdot)$ denote the expectations from the true (un-

[41] However, we also consider the results for other values of SR and γ. As discussed further below, higher Sharpe ratios and lower risk aversion coefficients will generally result in greater utility benefits.

[42] We note that, while a proportional change in the Sharpe ratio and the relative risk aversion coefficient would imply the exact same risk target SR/γ, the resulting utility equals $SR^2/(2\gamma)$ so the risk target alone is not a sufficient statistic for the utility in general.

known) risk model. Assuming that the investor uses model θ, to choose the position $x_t^\theta = 10\%/\sqrt{E_t^\theta(RV_{t+1})}$, the expected utility per unit of wealth, $UoW_t^\theta \equiv U_t(x_t^\theta)/W_t$, may be expressed as,

$$UoW_t^\theta = 3\% \frac{\sqrt{E_t(RV_{t+1})}}{\sqrt{E_t^\theta(RV_{t+1})}} - 1.5\% \frac{E_t(RV_{t+1})}{E_t^\theta(RV_{t+1})}. \quad (19)$$

We evaluate this expected utility empirically by averaging the corresponding realized expressions over the same rolling out-of-sample risk model forecasts underlying the results discussed in Section 5.2,

$$UoW^\theta = \frac{1}{T}\sum_{t=1}^{T}\left(3\% \frac{\sqrt{RV_{t+1}}}{\sqrt{E_t^\theta(RV_{t+1})}} - 1.5\% \frac{RV_{t+1}}{E_t^\theta(RV_{t+1})}\right). \quad (20)$$

For short, we will simply refer to this as the "realized utility." In parallel to (16), a risk model that perfectly predicts the realized volatilities, delivers a realized utility of $3\% - 1.5\% = 1.5\%$. Or said differently, the value of trading the risky asset with the perfect risk model is worth 1.5% percent of wealth. Importantly, the expected returns do not enter this expression. As such, this circumvents the invariable noise stemming from random return realizations, thereby allowing for a more pointed and meaningfully comparison of the different risk models based solely on the actual realized volatilities and the $E_t^\theta(RV_{t+1})$ risk model forecasts.

6.2. Realized Utility Comparisons

We begin by studying the utility benefits in the absence of transaction costs. Guided by the out-of-sample results in Section 5.2, we only consider the mega versions of the different risk models that restrict the estimated coefficients to be the same for all assets. The resulting realized utilities averaged across all assets, reported in the first row in Table 6, are obviously all close. Nonetheless, the corresponding DM-tests comparing the average realized utility for the HExp model to the realized utilities for each of the other models, reported in the first row in the lower panel of the table, show that the 1.30% utility obtained with HExp model is significantly higher than the 1.29% utility for the 21-Day RV, HAR and MIDAS models.[43] Further, consistent with the ordering of the different risk models based on the out-of-sample R^2's in Table 5, the HExpGl model results in significantly higher utility than the simple HExp model.

For comparison purposes, the last two columns give the realized utilities using the actual *future* realized monthly and daily volatilities, respectively. As previously noted, perfectly

[43] Even though the differences in the average realized utilities are numerically small, the HExp model almost always performs the best, thus explaining the significance of the t-statistics.

targeting the *daily* volatility would result in a realized utility of 1.50%. By comparison, a risk model that was able to perfectly predict the *monthly* volatilities would result in an average realized utility of 1.34% (which is lower than 1.50% because here the volatility is not exactly targeted on each day within the month). This number is surprisingly close to the utilities obtained from each of the practically feasible RV-based risk models, that is, our practically feasible models achieve almost the same utility of a "cheating" approach of exactly targeting the future monthly RV.

To put the results from the dynamic risk models further into perspective, the average realized utility from a static risk model based on the full in-sample unconditional volatility equals 0.75%. Hence, a risk-targeting investor would be willing to sacrifice 1.30% - 0.75% = 0.55% of his wealth annually to be able to use the benchmark HExp model instead of a static risk model. This 0.55% utility gain for the HExp model is obtained under the assumption of a Sharpe ratio of 0.3 and a risk aversion coefficient of 3, or equivalently from equation (15) an optimal risk target of 10% together with a Sharpe ratio of 0.3. Table 7 reports the utility gains of the HExp model that would result under alternative assumptions about the risk target and Sharpe ratio. Clearly, the utility benefits of the HExp (and other dynamic risk models) are substantially larger for larger SR's and/or risk targets. For instance, if the SR is 0.4 and the annualized risk target is 15% (corresponding to a risk aversion of 2.7), then the utility benefits double at 1.09% per year.

As discussed above, the benefits from risk targeting are also intimately related to the volatility-of-volatility: if the volatility was constant through time, the optimal positions would also be constant. The disaggregated results for the different asset classes reported in the Supplementary Appendix directly illustrates this. In particular, fixed income, which had the most stable volatility over the sample period, shows the smallest relative gains in utility from 0.90% for constant weights, to 1.30% for the HExp model. By comparison, for equities, which exhibited the highest asset class volatility-of-volatility over the sample, a constant investment resulted in a realized utility of only 0.57%, compared to 1.29% for the HExp risk model.

The previous comparisons ignore the cost of implementing the risk targeting positions. In actuality, of course, trading is costly. Since the realized utility in (20) is effectively expressed in units of returns, it is easy to incorporate the effect of transaction costs by simply subtracting the simulated costs of implementing the positions. For simplicity, we assume that the costs of trading are linear in the absolute magnitude of the change in the positions,

$|x^{\theta}_{t-1} - x^{\theta}_{t}|$.[44] As our benchmark transaction cost estimate, we use one-half times the median bid-ask spread for each of the assets over the last nine months of the sample.[45] Said differently, the transaction cost is the difference between the mid-quote and the bid or ask price. The transaction costs could be higher for a large trader due to market impact. On the other hand, the transaction costs could be lower due to the possibility of strategic trading and transactions occurring inside the spread combined with the netting of other positions.[46]

The second row in Table 6 reports the resulting realized utilities net of transaction costs. All of the utilities are obviously lower than the ones reported in the first row. Interestingly, the perfect-foresight one-day-ahead volatility positions reported in the last column is now quite inferior and far from the 1.50% maximum utility, as the cost of (too much) trading far outstrip the benefits of continuously adjusting the optimal positions. Similarly, the benefits of perfectly knowing the true monthly volatilities are also substantially reduced compared to the situation with "free" trading. Comparing the different risk models, the DM-tests reported in the lower panel of the table, show that even though numerically close, the realized utility for the MIDAS model is now significantly higher than the utility for the simple HExp model. This change in the ordering of the models is readily explained by the previous Figure 6, and the "slower" response of the MIDAS model to more recent information, which effectively serves to diminish the change in the optimal positions thereby reducing transaction costs.

A commonly employed approach to help mitigate the impact of transaction costs more generally is to slow down trading, allowing the investor to deviate from the zero-cost optimal positions. The formal development of optimal trading strategies for the different risk models that explicitly incorporate transaction costs is beyond the scope of this paper. Instead, we rely on the strategy discussed by (Garleanu and Pedersen, 2013, 2016) of trading only partially toward the desired position.[47] The third row in Table 6, in particular, reports the realized utilities when the "speed" of trading is reduced, and the positions are only traded half-way toward the zero-cost optimal targets. This "slowing down" of the models does indeed restore some of the "lost" utility vis-a-vis the results in the first row that abstracts from transaction costs and the results in the second row based on trading all the way to the targeted positions. The HExp and HExpGl models now also again significantly outperform

[44]In practice, of course, the cost will generally increase with the magnitude of the trade, and will likely also depend on the volatility. We leave a more thorough analysis of these issues for future work.

[45]The only exception is USDSEK (Sweden FX), where we use spread data from all of 2013 due to data availability.

[46]Additional results for other transaction cost assumptions are reported in the Supplementary Appendix.

[47](Garleanu and Pedersen, 2013, 2016) show that this strategy is optimal under certain conditions, including quadratic transaction costs due to, e.g., market impact. Thus, even though this strategy is not fully optimal here, it nevertheless gives a sense of the benefits that may obtain by reducing the "speed" of trading.

all of the other practically feasible risk models, although the differences in the actual realized utilities are fairly small.

7. Conclusion

This paper studies risk across commodities, currencies, global equities, and fixed income securities using a new extensive dataset of intraday data. Based on this rich dataset, we find that risk dynamics are surprisingly similar across assets, asset classes, and countries. Normalized by the average levels, risk measures everywhere have similar unconditional distributions and autocorrelation structures. This commonality in risk structures can be exploited when estimating risk models by aggregating information across assets using panel estimation methods. Beyond common structures in asset risk levels and dynamics, we also find that a common normalized "global" risk factor contains information on the future volatilities of individual assets not already contained in the asset specific realized volatility histories, and that models which include this common factor outperform models that do not.

In addition to the robustness generated by exploiting commonality in risk dynamics, we add further robustness by developing new "smooth" and "centered" realized volatility models that enforce a natural continuous and monotonic dependence on lagged realized volatilities. We show that these new robust risk models perform well in out-of-sample risk forecasting.

Lastly, we develop a simple framework for quantifying the utility benefits of risk models for risk-targeting investors. We show that under empirically realistic assumptions, our robust dynamic risk models are worth at least 0.55% of wealth per year relative to a static risk model.

8. References

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., Diebold, F. X., 2006. Volatility and correlation forecasting. In: Handbook of Economic Forecasting (eds. G. Elliott, C.W.J. Granger and A. Timmermann).

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., Diebold, F. X., 2013. Financial risk measuremet for financial risk management. In: Handbook of the Economics of Finance (eds. G. Constantinides, M. Harris and R. Stulz).

Andersen, T. G., Bollerslev, T., Diebold, F. X., 2007. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. Review of Economics and Statistics 89, 701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2000. Great realizations. Risk, 105–108.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2003a. Modeling and forecasting realized volatility. Econometrica 71, 579–625.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Vega, C., 2003b. Micro effects of macro announcements: Real-time price discovery in foreign exchange. American Economic Review 93, 38–62.

Andersen, T. G., Bollerslev, T., Meddahi, N., 2011. Realized volatility forecasting and market microstructure noise. Journal of Econometrics 160, 220–234.

Audrino, F., Knaus, S., 2016. Lassoing the HAR model: A model selection perspective on realized volatility dynamics. Econometric Reviews 35, 1485–1521.

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

Baker, M., Wurgler, J., Yuan, Y., 2012. Global, local, and contagious sentiment. Journal of Financial Economics 104, 272–287.

Bakshi, G., Madan, D., 2006. A theory of volatility spreads. Management Science 52, 1945–1956.

Balduzzi, P., Lynch, A. W., 1999. Transaction costs and predictability: Some utility cost calculations. Journal of Finance 52, 47–78.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2008. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. Econometrica 76, 1481–1536.

Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. Review of Financial Studies 22, 4463–4492.

Boudoukh, J., Feldman, R., Kogan, S., Richardson, M., 2016. Information, trading, and volatilty: Evidence from firm-specific news. Working Paper, Hebrew University, University of Texas at Austin, and New York University.

Brown, G. W., 1999. Volatility, sentiment, and noise traders. Financial Analyst Journal 55, 82–90.

Campbell, J. Y., Thompson, S. B., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? Review of Financial Studies 21, 1509–1531.

Chen, X., Ghysels, E., 2011. News – good or bad – and its impact on volatility predictions over multiple horizons. Review of Financial Studies 24, 46–81.

Clements, M. P., Galvao, A. B., Kim, J. H., 2008. Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. Journal of Empirical Finance 15, 729–750.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics 7, 174–196.

Craioveanu, M., Hillebrand, E., 2012. Why is it ok to use the HAR-RV(1,5,22) model? Working Paper, CREATES, University of Aarhus, Denmark.

Diebold, F., 2015. Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold-Mariano tests. Journal of Business and Economic Statistics 33, 1–9.

Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253–263.

Ederington, L. H., Lee, J. H., 1993. How markets process information: News releases and volatility. Journal of Finance 48, 1161–1191.

Engle, R., Ito, T., Lin, W., 1990. Meteor showers of heat waves? Heteroskedastic intradaily volatiloty in the foregn exchange market. Econometrica 58, 525–542.

Engle, R., Mezrich, J., 1996. GARCH for groups. Risk 9, 36–40.

Engle, R., Rangel, J., 2008. The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. Review of Financial Studies 21, 1187–1222.

Fleming, J., Kirby, C., Ostdiek, B., 2001. The economic value of volatility timing. Journal of Finance 56, 329–352.

Fleming, J., Kirby, C., Ostdiek, B., 2003. The economic value of volatility timing using realized volatility. Journal of Financial Economics 67, 473–509.

Foroni, C., Marcellino, M., Schumacher, C., 2015. Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. Journal of the Royal Statistical Soceity 178, 57–82.

French, K., Roll, R., 1986. Stock return variances: The arrival of information and the reaction of traders. Journal of Financial Economics 17, 5–26.

French, K. R., Schwert, G. W., Stambaugh, R. F., 1987. Expected stock returns and volatility. Journal of Financial Economics 19, 3–30.

Gallant, R. A., Hsu, C. T., Tauchen, G., 1999. Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. Review of Economics and Statistics 81, 617–631.

Garleanu, N., Pedersen, L. H., 2013. Dynamic trading with predictably returns and transaction costs. Journal of Finance 68, 2309–2340.

Garleanu, N., Pedersen, L. H., 2016. Dynamic portfolios choice with frictions. Journal of Economic

Theory 165, 487–516.

Ghysels, E., Qian, H., 2016. Estimating MIDAS regressins via ols with polynomial parameter profiling. Working Paper, UNC.

Ghysels, E., Rubia, A., Valkanov, R., 2009. Multi-period forecasts of variance: Direct, iterated, and mixed-data approaches. Working Paper, Alicante University, UCSD and UNC.

Ghysels, E., Santa-Clara, P., Valkanov, R., 2006. Predicting volatility: getting the most out of return data sampled at different frequencies. Journal of Econometrics 131, 59–96.

Ghysels, E., Sinko, A., 2011. Volatility forecasting and microstructure noise. Journal of Econometrics 160, 257–271.

Ghysels, E., Sinko, A., Valkanov, R., 2007. MIDAS regressions: further results and new directions. Econometric Reviews 26, 53–90.

Hansen, L. P., Hodrick, R. J., 1980. Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. Journal of Political Economy 88, 829–853.

Hansen, P. R., Lunde, A., 2005a. A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? Journal of Applied Econometrics 20, 873–889.

Hansen, P. R., Lunde, A., 2005b. A realized variance for the whole day based on intermittent high-frequency data. Journal of Financial Econometrics 3, 525–554.

Hansen, P. R., Lunde, A., 2014. Estimating the persistence and the autocorrelation function of a time series that is measured with error. Econometric Theory 30, 60–93.

Hastie, T., Tibshirani, R., Friedman, J., 2009. Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer Verlag.

Hsieh, D. A., 1991. Chaos and nonlinear dynamics: Application to financial markets. Journal of Finance 46, 1839–1877.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., Vetter, M., 2009. Microstructure noise in the continuous case: The pre-averaging approach. Stochastic Processes and their Applications 119, 2249–2276.

Karolyi, A. G., 1993. A Bayesian approach to modeling stock return volatility for optin valuation. Journal of Financial and Quantitative Analysis 28, 579–593.

Karolyi, A. G., 1995. A multivariate GARCH model of international transmissions of stock returns and volatility: The case of the United States and Canada. Journal of Business and Economic Statistics 13, 11–25.

Karolyi, G., Lee, K., van Dijk, M., 2012. Understanding commonality in liquidity around the world. Journal of Financial Economics 105, 82–112.

Karpoff, J. M., 1987. The relation between price changes and trading volume: A survey. Journal of Financial and Quantitative Analysis 22 (1), 109–126.

Liu, L., Patton, A. J., Sheppard, K., 2015. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. Journal of Econometrics 187, 293–311.

Lynch, A. W., Tan, S., 2010. Multiple risky assets, transaction costs, and return predictability: Allocation rules and implications for us investors. Journal of Financial and Quantitative Analysis, 1015–1053.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–08.

Patton, A. J., Sheppard, K., 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. Review of Economics and Statistics 97, 683–697.

Pedersen, L. H., 2015. Efficiently Inefficient: How Smart Money Invests and Market Prices are Determined. Princeton University Press.

Sangvinatsos, A., Wachter, J. A., 2005. Does the failure of the expectations hypothesis matter for long-term investors? Journal of Finance 60, 179–230.

Schwert, G. W., 1989. Why does stock market volatility change over time? Journal of Finance 44, 1115–1153.

Sizova, N., 2011. Integrated variance forecsting: Model based vs. reduced form. Journal of Econometrics 162, 294–311.

Swanson, N. R., White, H., 1997. Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models. International Journal of Forecasting 13, 439–461.

Taylor, S., 1986. Modelling Financial Time Series. John Wiley and Sons.

Taylor, S., 2005. Asset Price Dynamics, Volatility, and Prediction. Princeton University Press.

Yu, J., Yuan, Y., 2011. Investor sentiment and the mean-variance relation. Journal of Financial Economics 100, 367–381.

Zhang, L., Mykland, P. A., Aït-Sahalia, Y., 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association 100, 1394–1411.

Tables

Table 1: *RV* **Summary Statistics**

	Commodities	Equities	Fixed Income	Foreign Exchange
Mean	25.4	20.6	3.1	10.3
St.dev.	12.6	13.7	1.5	5.7
Skewness	2.6	3.4	2.3	3.1
Excess kurtosis	16.9	22.9	11.6	18.5
Maximum	185.6	186.6	19.4	74.1
95th percentile	47.8	44.8	5.8	20.4
50th percentile	22.7	17.0	2.8	9.0
5th percentile	11.6	8.2	1.5	4.6
Minimum	4.9	3.0	0.6	1.2
1-Day autocorr.	0.516	0.707	0.481	0.517
20-Day autocorr.	0.362	0.480	0.347	0.415
100-Day autocorr.	0.195	0.228	0.197	0.221
250-Day autocorr.	0.115	0.105	0.073	0.104
Number of assets	20	21	8	9
Ave. number of obs.	5407	3812	4042	3351
Earliest start date	10/22/1992	1/3/1996	1/3/1996	1/1/1999
Latest start date	1/3/1996	12/13/2005	9/26/2000	1/1/2004

Notes: The table presents summary statistics of daily realized volatilities averaged across all assets within a given asset class. All of the volatility numbers are reported in annualized percentage units.

Table 2: RV **Contemporaneous and Partial Correlations**

	Commodities	Equities	Fixed Income	Foreign Exchange
Panel A: Contemporaneous correlations				
Commodities	0.277	0.298	0.217	0.355
Equities		0.668	0.407	0.554
Fixed Income			0.470	0.433
Foreign Exchange				0.710
Global	0.410	0.721	0.547	0.633
Panel B: Partial correlations and ΔR^2				
Commodities	0.059	0.063	0.074	0.132
Equities	0.081	0.173	0.122	0.198
Fixed Income	0.072	0.078	0.093	0.151
Foreign Exchange	0.078	0.108	0.110	0.144
Global	0.171	0.328	0.267	0.456
Commodities	0.005	0.004	0.007	0.017
Equities	0.011	0.021	0.024	0.041
Fixed Income	0.006	0.005	0.011	0.016
Foreign Exchange	0.007	0.010	0.015	0.014
Global	0.012	0.022	0.033	0.061

Notes: The table presents correlations of daily realized volatilities averaged within and across asset classes. The top Panel A reports the standard contemporaneous correlations. The bottom Panel B reports the average partial correlation coefficients obtained from regressions of the daily RVs for the assets within the asset classes indicated in the columns on their own daily lags and the lagged values of the RVs for the assets in the asset classes indicated in the rows (top panel), together with the average absolute percentage increases in the R^2s compared to regressions of the assets indicated in the columns on their own daily lags only (bottom panel). The "Global" volatility factor is constructed as a weighted average of the daily RVs, as further discussed in the main text.

Table 3: **Global Volatility Correlations**

	Sentiment	Unusual Sentiment	VRP	News
Global Volatility	-0.123*	0.197**	-0.454**	0.183**
Sentiment		0.568**	-0.097	0.171**
Unusual Sentiment			0.051	0.136*
VRP				-0.087

Notes: The table reports the monthly (end-of-month) contemporaneous correlations between the global volatility factor *ExpGlRV* formally defined in Section 4.6 and four other variables. Sentiment refer to the "orthogonalized" investor sentiment measure of Baker and Wurgler (2006). Unusual Sentiment is defnied as the absolute value of the Sentiment measure minus its sample mean. The variance risk premium (VRP) is equal to the difference between the VIX and the realized U.S. equity volatility over the past month. The News surprise variable is constructed as the average value of the standardized absolute news announcement surprises observed over the month. Statistical significance at the 5% and 1% level are indicated by * and **, respectively.

Table 4: **In-Sample Predictions**

		21-Day RV	HAR	HAR-Free	MIDAS	HExp	HExpGl
				R^2			
Commodities	Indiv.	0.271	0.453	0.458	0.461	0.459	0.471
	Panel	0.271	0.434	0.437	0.436	0.438	0.448
	Mega	0.271	0.432	0.435	0.435	0.437	0.444
Equities	Indiv.	0.209	0.436	0.437	0.434	0.438	0.469
	Panel	0.209	0.421	0.420	0.416	0.422	0.452
	Mega	0.209	0.419	0.418	0.415	0.420	0.441
Fixed Income	Indiv.	0.276	0.431	0.449	0.441	0.467	0.474
	Panel	0.276	0.425	0.440	0.434	0.459	0.461
	Mega	0.276	0.416	0.428	0.422	0.439	0.423
Foreign Exchange	Indiv.	0.399	0.537	0.537	0.543	0.544	0.611
	Panel	0.399	0.530	0.529	0.535	0.535	0.588
	Mega	0.399	0.524	0.519	0.527	0.526	0.562
All Assets	Indiv.	0.242	0.445	0.448	0.448	0.449	0.471
	Panel	0.242	0.428	0.429	0.427	0.431	0.451
	Mega	0.242	0.426	0.427	0.426	0.429	0.443
				DM t-tests			
Commodities	Indiv.	-4.60	-2.22	-1.30	-0.46	NA	1.68
	Panel	-4.60	-2.45	-2.19	-2.22	-2.00	-1.27
	Mega	-4.60	-2.55	-2.31	-2.32	-2.11	-1.64
Equities	Indiv.	-2.40	-0.15	-0.63	-0.65	NA	1.29
	Panel	-2.40	-1.30	-1.50	-1.32	-1.41	0.28
	Mega	-2.40	-1.31	-1.54	-1.36	-1.44	-0.36
Fixed Income	Indiv.	-3.46	-2.10	-2.77	-1.42	NA	1.44
	Panel	-3.46	-2.30	-3.31	-1.75	-2.65	-1.92
	Mega	-3.46	-3.10	-4.24	-2.84	-3.66	-2.07
Foreign Exchange	Indiv.	-1.28	-0.60	-0.69	-0.15	NA	1.08
	Panel	-1.28	-1.20	-1.52	-0.97	-1.79	0.66
	Mega	-1.28	-2.01	-2.06	-1.79	-1.54	0.98
All Assets	Indiv.	-3.64	-0.86	-0.83	-0.23	NA	1.39
	Panel	-3.64	-2.20	-2.33	-2.02	-2.20	0.17
	Mega	-3.64	-2.24	-2.36	-2.10	-2.21	-0.65

Notes: The top panel reports the average in-sample regression R^2s for different risk models and estimation procedures by asset class and across all assets. The bottom panel reports Diebold-Mariano (DM) t-statistics for testing the significance of the average standardized loss differentials relative to the individually estimated HExp model. A positive (negative) DM-test statistic indicates that the model and estimation procedure outperforms (underperforms) the individually estimated HExp model in-sample.

Table 5: **Out-of-Sample Predictions**

	21-Day RV	HAR	HAR-Free	MIDAS	HExp	HExpGl
			R^2			
Indiv.	0.277	0.428	0.461	0.434	0.473	0.495
Panel	0.277	0.455	0.485	0.459	0.487	0.507
Mega	0.277	0.462	0.488	0.467	0.492	0.506
			DM t-tests			
Indiv.	-2.56	-2.01	-2.30	-2.02	-1.93	-0.23
Panel	-2.56	-2.83	-2.29	-3.43	-2.05	0.91
Mega	-2.56	-3.64	-1.95	-4.80	NA	1.21

Notes: The top panel reports the out-of-sample predictive R^2s for different risk models and estimation procedures averaged across all assets. The bottom panel reports Diebold-Mariano (DM) t-statistics for testing the significance of the average standardized loss differentials relative to the mega HExp model that restricts the coefficients to be the same across all assets. A positive (negative) DM-test statistic indicates that the model and estimation procedure outperforms (underperforms) the mega HExp model out-of-sample.

Table 6: **Realized Utility**

	Risk Model						Future RV	
Tcost/Speed	21-Day	HAR	HAR-Free	MIDAS	HExp	HExpGl	20-Day	1-Day
				Utility				
0/100%	1.29%	1.29%	1.30%	1.29%	1.30%	1.30%	1.34%	1.50%
0.5/100%	1.24%	1.23%	1.24%	1.25%	1.25%	1.25%	1.29%	0.61%
0.5/50%	1.25%	1.25%	1.26%	1.26%	1.26%	1.27%	1.31%	1.06%
				DM t-tests				
0/100%	-2.96	-9.50	0.69	-10.86	NA	3.78	11.71	36.10
0.5/100%	-1.92	-9.75	-2.03	5.00	NA	2.04	11.32	-30.12
0.5/50%	-4.24	-7.57	-0.81	-3.97	NA	2.23	10.50	-18.75

Notes: The top panel reports the utilities from holding volatility-targeted positions averaged across all of the assets. The first column gives the transactions costs in multiples of the median spread for each of the assets, together with the "speed" of the trading rule based on how close to the targeted position the assets are traded at the close of each business day. All of the results are based on rolling out-of-sample predictions from mega-based estimation of the different risk models that restrict the coefficients to be the same across all assets. The results listed under "Future RV" rely on the actual realized future RVs for determining the positions. The bottom panel reports Diebold-Mariano (DM) t-statistics for testing the significance of the average utility differentials relative to the HExp model under the identical transaction cost and trading speed assumptions. A positive (negative) DM-test statistic indicates that the alternative procedure outperforms (underperforms) the mega HExp model in terms of the out-of-sample realized utility gains.

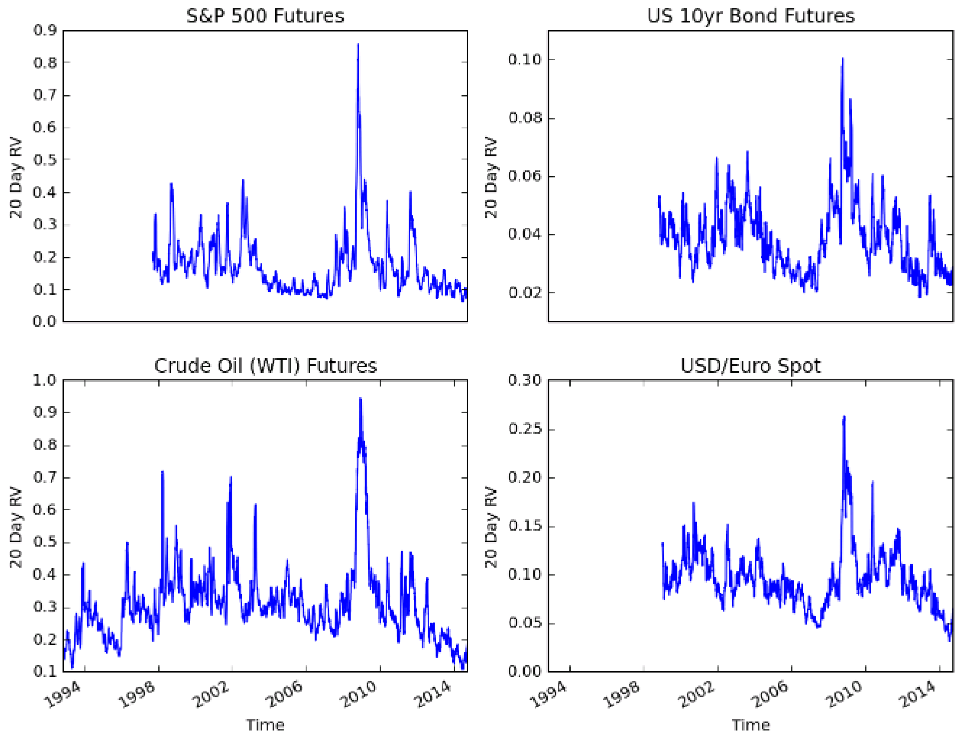
Table 7: **Utility Benefits from Dynamic Risk Modeling**

	Risk target				
SR	5%	10%	15%	20%	25%
0.1	0.09%	0.18%	0.27%	0.36%	0.45%
0.2	0.18%	0.36%	0.55%	0.73%	0.91%
0.3	0.27%	0.55%	0.82%	1.09%	1.36%
0.4	0.36%	0.73%	1.09%	1.46%	1.82%
0.5	0.45%	0.91%	1.36%	1.82%	2.27%

Notes: The table reports the average utility benefits from optimal positions predicted by the HExp risk model relative to a static risk model for different Sharpe Ratios (SR) and risk targets. The calculations exclude transactions costs and assume that all assets are trade all the way to their targeted positions at the close of each business day. The results reported in Table 6 are based on a Sharpe ratio of 0.3 and a risk target of 10%, or equivalently a Sharpe ratio of 0.2 and a risk target of 15%, corresponding to an average gain of 0.55%.

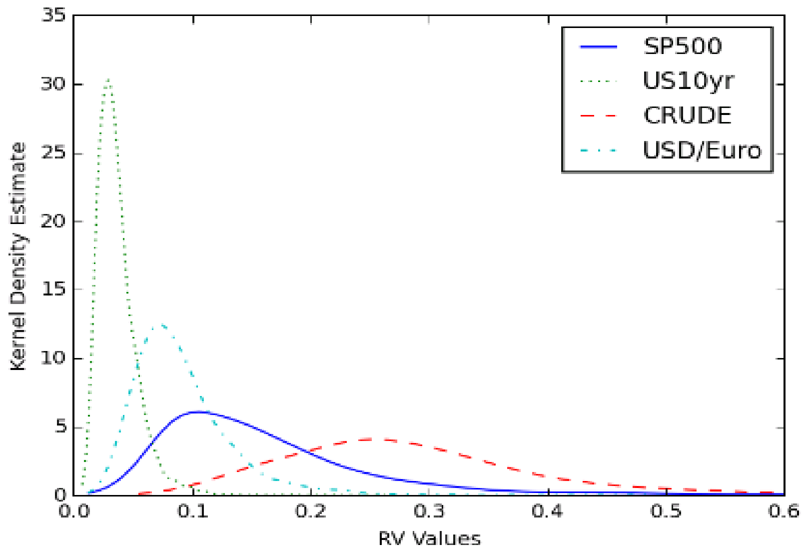
Figures

Figure 1: **Monthly Realized Volatilities**



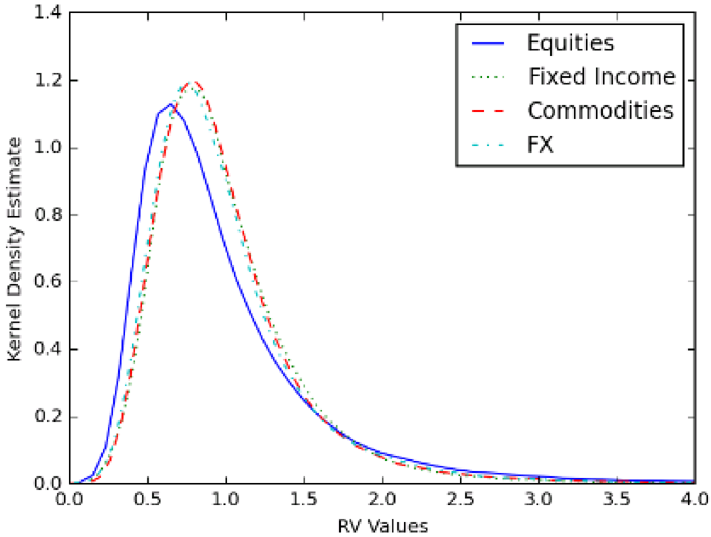
Note: This figure shows the timeseries of 20-day average realized volatilities (annualized) for four representative assets: S&P 500 Futures, US 10yr Bond Futures, Crude Oil (WTI) Futures, and USD/Euro Spot.

Figure 2: **Unconditional Daily *RV* Distributions**



Note: This figure shows kernel density estimates of the unconditional daily realized volatility estimates (annualized) for four representative assets: S&P 500 Futures, US 10yr Bond Futures, Crude Oil (WTI) Futures, and USD/Euro Spot.

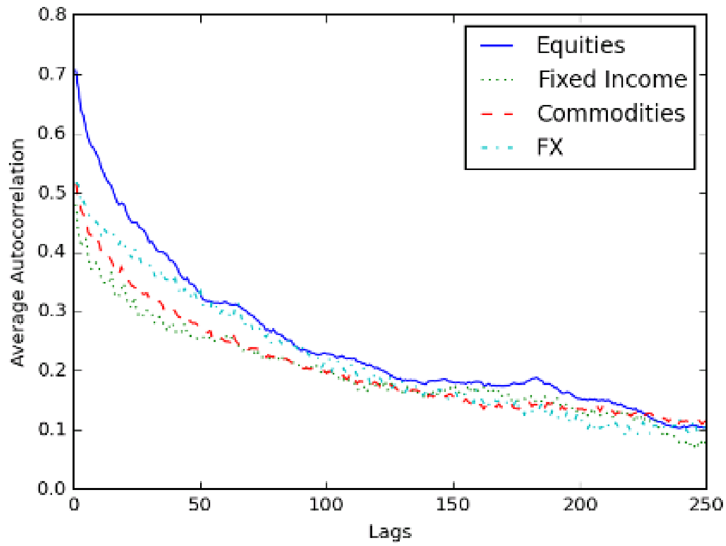
Figure 3: **Normalized Unconditional Daily RV Distributions Distributions**



Note: This figure shows kernel density estimates of daily realized volatility distributions (annualized) of all assets within each of the four asset classes, where each asset's daily RV is divided by its own asset's average daily RV ("Normalized") to account for level differences in individual asset volatilities.

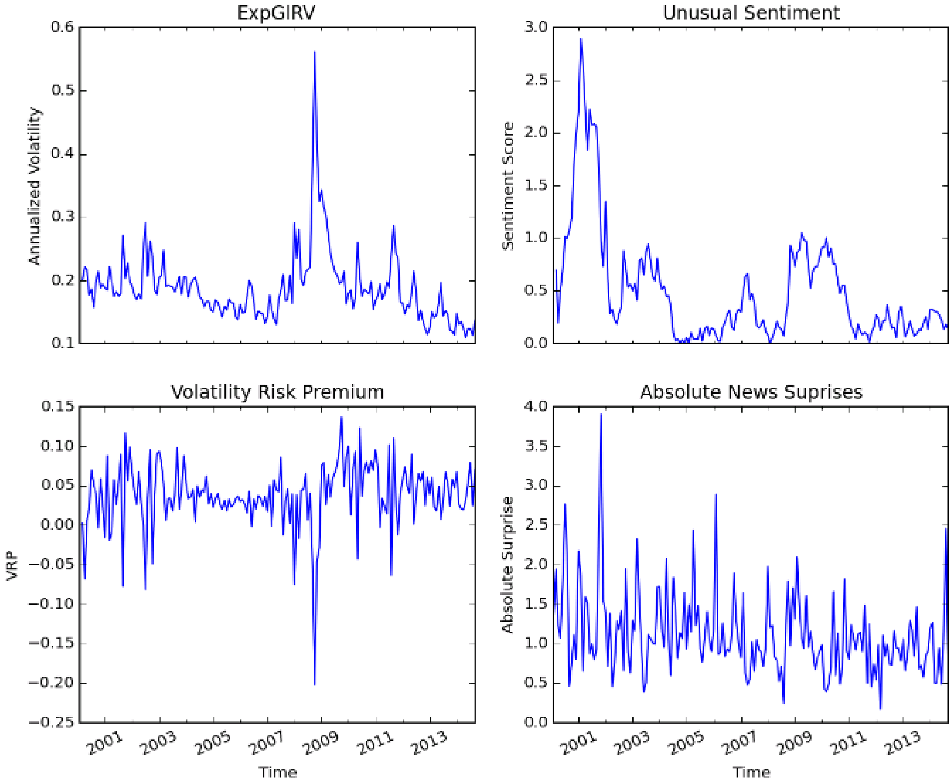
Figure 4: **Daily Realized Volatility Autocorrelations**

Figure 4: Daily Realized Volatility Autocorrelations



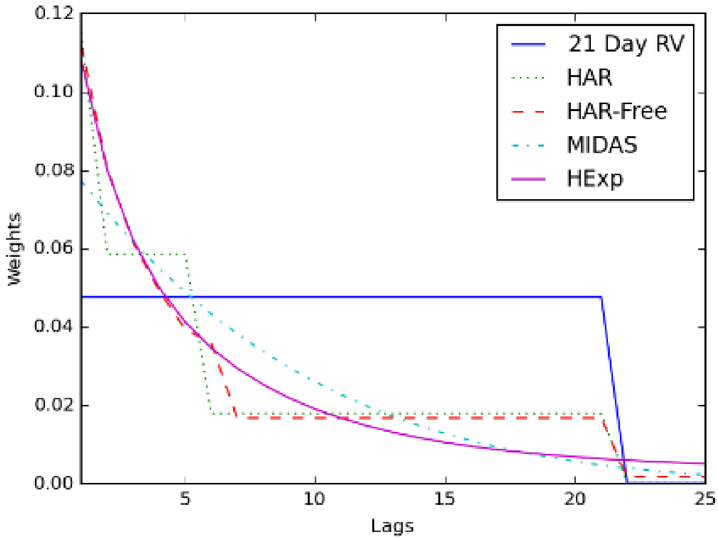
Note: This figure shows the average autocorrelation function of the daily realized volatilities averaged across all of the assets within each of the four asset classes.

Figure 5: **Global Volatility, Sentiment, VRP, and News**



Note: This figure shows the global volatility factor (ExpGIRV), unusual monthly sentiment, the volatility risk premium (VRP), and the absolute news surprise variable, as formally defined in the main text. All of the variables are plotted at a monthly frequency.

Figure 6: **Implied Lag Coefficients for Different Risk Models**



Note: This figure shows the lag coefficients implied by the regression coefficients of full period models pooled across all assets for each of five different RV-based models: 21 Day RV, HAR, HAR-Free, MIDAS, and HExp.

Appendix

A.1. Data and Data Cleaning

Our data for the different assets come from a few different data sources. The data starts at different points in time, but all end on September 30, 2014. Table A.1 provides a summary.

Asset Class	Asset	Number of Assets	Total Days in Analysis	Primary Data Source	Used From	Secondary Data Source	Used From	Assumed T-Costs (in bps)
COMMODITIES		**20**	**108149**	**TRTH**		**TDC**		
	Brent Oil	1	4754	TRTH	1/3/1996	TDC	1/3/1996	1.0
	Cattle	1	5483	TRTH	12/20/2004	TDC	11/30/1992	3.2
	Cocoa	1	5471	TRTH	4/1/2008	TDC	11/11/1992	3.4
	Coffee	1	5469	TRTH	4/1/2008	TDC	11/17/1992	8.0
	Corn	1	5502	TRTH	8/1/2006	TDC	11/19/1992	5.7
	Cotton	1	5453	TRTH	4/1/2008	TDC	11/12/1992	4.6
	Crude (WTI) Oil	1	5480	TRTH	9/5/2006	TDC	11/10/1992	1.0
	Feeder Cattle	1	5513	TRTH	8/1/2007	TDC	10/29/1992	4.5
	Gas Oil	1	4754	TRTH	1/3/1996	TDC	1/3/1996	2.8
	Gold	1	5471	TRTH	12/4/2006	TDC	12/2/1992	0.8
	Heating Oil	1	5480	TRTH	9/5/2006	TDC	11/16/1992	1.7
	Lean Hogs	1	5486	TRTH	2/15/2005	TDC	11/30/1992	4.5
	Natural Gass	1	5442	TRTH	8/23/2006	TDC	1/5/1993	4.0
	Silver	1	5412	TRTH	12/4/2006	TDC	1/5/1993	2.6
	Soybeans	1	5522	TRTH	8/1/2006	TDC	10/22/1992	2.1
	Soymeal	1	5502	TRTH	8/1/2006	TDC	11/19/1992	4.1
	Soyoil	1	5501	TRTH	8/1/2006	TDC	11/19/1992	3.0
	Sugar	1	5481	TRTH	4/1/2008	TDC	11/3/1992	5.9
	Unleaded (RBOB)	1	5475	TRTH	8/22/2006	TDC	11/16/1992	2.0
	Wheat	1	5498	TRTH	8/1/2006	TDC	11/19/1992	4.4
EQUITIES		**21**	**80042**	**TRTH**		**NONE**		
	Australia (SPI 200)	1	3472	TRTH	12/18/2000	NA	NA	1.9
	Germany (DAX 30)	1	4732	TRTH	1/3/1996	NA	NA	1.0
	Brazil (BOVESPA)	1	4577	TRTH	2/27/1996	NA	NA	2.8
	China (Hang Seng CEI)	1	2667	TRTH	12/9/2003	NA	NA	2.0
	Canada (S&P/TSX 60)	1	3773	TRTH	9/14/1999	NA	NA	1.3
	Spain (IBEX 35)	1	4698	TRTH	1/4/1996	NA	NA	2.0
	Eurostoxx	1	4130	TRTH	6/23/1998	NA	NA	3.2
	France (CAC 40)	1	4007	TRTH	1/7/1999	NA	NA	1.1
	Hong Kong (Hang Seng)	1	4591	TRTH	1/3/1996	NA	NA	1.2
	India (SGX NIFTY)	1	2213	TRTH	10/11/2005	NA	NA	1.7
	Italy (FTSE MIB)	1	2617	TRTH	6/15/2004	NA	NA	2.4
	Japan (TOPIX)	1	4570	TRTH	1/5/1996	NA	NA	4.1
	South Korea (KOSPI 200)	1	4466	TRTH	5/6/1996	NA	NA	1.9
	Netherlands (AEX)	1	4499	TRTH	1/9/1997	NA	NA	1.3
	South Africa (ALSI)	1	2308	TRTH	7/7/2005	NA	NA	1.7
	Switzerland (SMI)	1	4027	TRTH	9/15/1998	NA	NA	1.2
	Taiwan (SGX-MSCI Taiwan)	1	4295	TRTH	2/24/1997	NA	NA	3.1
	UK (FTSE 100)	1	4706	TRTH	1/3/1996	NA	NA	0.8
	US (S&P 500 E-Mini)	1	4274	TRTH	9/10/1997	NA	NA	1.3
	US (Russell 2000 E-Mini)	1	2234	TRTH	12/13/2005	NA	NA	0.9
	US (S&P 400 Mid Cap E-Mini)	1	3186	TRTH	1/29/2002	NA	NA	1.5
FIXED INCOME		**8**	**32333**	**TRTH**		**TDC**		
	Australia 10y	1	4734	TRTH	1/3/1996	NA	NA	3.9
	Germany 10y	1	4499	TRTH	1/5/1999	TDC	1/3/1997	0.7
	Germany 5y	1	4493	TRTH	2/1/1999	TDC	1/3/1997	0.8
	Canada 10y	1	2771	TRTH	9/26/2000	NA	NA	0.8
	Japan 10y	1	3605	TRTH	1/5/1996	NA	NA	0.7
	UK 10y	1	4711	TRTH	1/3/1996	NA	NA	0.9
	US 10y	1	3993	TRTH	1/1/2001	TDC	10/20/1998	1.3
	US 5y	1	3527	TRTH	7/1/2001	TDC	9/5/2000	0.7
FOREIGN EXCHANGE		**9**	**30161**	**Olsen Data**		**NONE**		
	Australia (AUD-USD)	1	2802	OlsenData	1/1/2004	NA	NA	2.2
	Eurozone (EUR-USD)	1	4103	OlsenData	1/1/1999	NA	NA	0.7
	Canada (USD-CAD)	1	3061	OlsenData	1/1/2003	NA	NA	2.5
	Japan (USD-JPY)	1	3841	OlsenData	1/1/2000	NA	NA	1.0
	Norway (USD-NOK)	1	2801	OlsenData	1/1/2004	NA	NA	8.1
	New Zealand (NZD-USD)	1	2803	OlsenData	1/1/2004	NA	NA	4.7
	Sweden (USD-SEK)	1	3062	OlsenData	1/1/2003	NA	NA	7.7
	Switzerland (USD-CHF)	1	3844	OlsenData	1/1/2000	NA	NA	1.7
	UK (GBP-USD)	1	3844	OlsenData	1/1/2000	NA	NA	1.5

A.1.1. Contract Rolling

Our returns for commodities, equities, and fixed income are all constructed from futures contract prices. Unlike spot prices (like our foreign exchange data), individual futures contracts are listed and subsequently expire multiple times per year.

With multiple futures contracts for the same underlying asset trading at the same time, we need a rule for determining which contracts to actually use on any given day. In practice, an asset manager typically gains continuous exposure through these contracts by "rolling" their position from a contract that is set to expire to the next contract in line. For instance, if S&P 500 E-mini futures come in contracts that expire in mid-March and mid-June, then an investor can hold the March-expiring contract through the beginning of March, and then close that position while simultaneously opening a comparable position in the E-mini contract that does not expire until June. Naturally, many investors would like to roll their positions at the same time.[48] Rolling can typically be done at far lower cost and market impact than outright trading for this reason, assuming that the trader rolls positions at the same time that many investors are seeking to roll their positions as well.

The dates when most investors roll positions (and therefore the times when the "roll trade" has the most liquidity) vary for different assets, but typically fall around 3-6 days prior to contract expiry or first notice date for non-commodities. Commodity roll times are more likely to be associated with the roll schedule of major commodity indices, such as the S&P GSCI. The fact that most of our assets studied are futures contracts that need to be rolled necessitates a formal "roll rule."[49] As an actual investor would, we determine our roll time based on a measure of liquidity in the near and far contracts. For each contract, we specify a roll period (typically 3-6 days prior to expiry for non-commodities, and the roll period for the GSCI for commodities), measuring the total number of minutes with at least one valid trade in them for both the near and far contract. Once the number of valid minute bars for the far contract exceeds the near contract over a trading day, or we reach the end of the roll period, we consider that day to be the roll date.[50]

[48] Recall that all futures contract positions have offsetting positions on the other side of the contract, so if investor A is long one contract and investor B is short one contract and they both desire to keep those positions, they can mutually close their position in the expiring contract (the "near" contract) and open a new position in the contract with the next expiry (the "far" contract) without requiring outright liquidity from the market.

[49] For example, rather than having a single asset representing the S&P 500 index, we have data for a series of different S&P 500 index futures contracts and a roll rule that we specify to connect them into a single return series.

[50] Since we only seek to compute daily RV's, it is possible to determine the occurrence of a roll date ex post and splice return sequences from both the near and far contracts together at the end of the trading day.

Once we have a roll date, we seek to line up the two contracts before switching our reference from "near" to "far." Omitting the last five minutes of the trading day to avoid any possible end-of-day effects, we find the last (nearest to the close) minute bar where both contracts have at least one trade in order to switch between contracts. This ensures that both contracts have non-stale prices at the time of the change. We use all returns from the near contract up to and including the roll minute bar, and then use returns from the far contract after that.

A.1.2. Liquidity Plots and "Sanity" Filters

To begin, we use the Financial Calendars (FinCal) to provide market open and close times for some of the market sessions dating back to the year 2000. For assets with data prior to 2000, or assets outside the FinCal database, we rely on so-called "liquidity plots," in which for each minute of the day, we plot the proportion with at least one trade. As the resulting liquidity plots for the four representative assets and three years given in Figure A.1 show, this effectively delineate the periods of the day when the markets are actively operating.

We also apply some mild "sanity" filters to identify "valid" minute bars. To begin, all valid minute bars must have at least one recorded trade over the course of the minute. In addition, we omit data when:

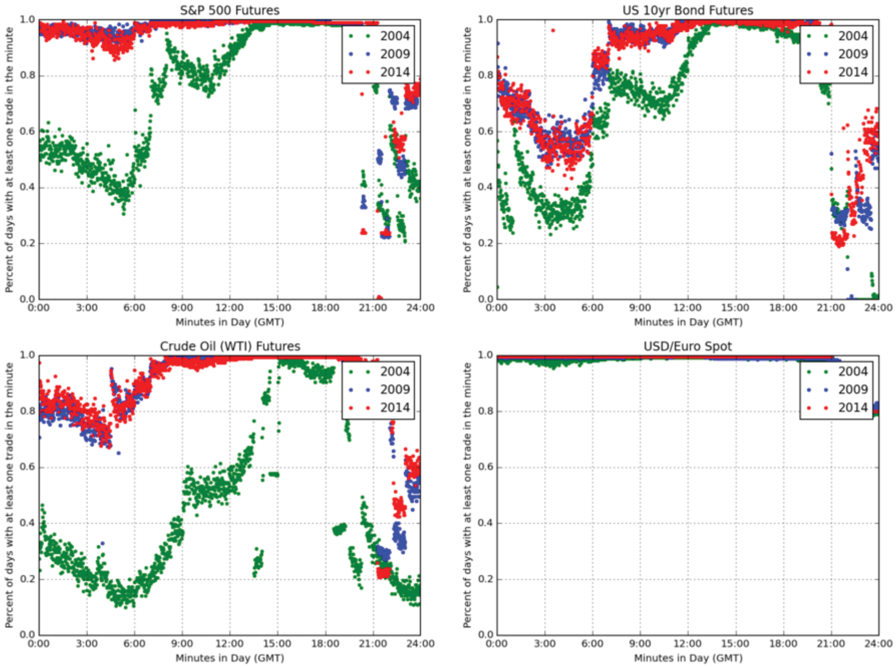
- The bid or the ask price is less than zero
- The ask price is less than or equal to the bid price
- The bid-ask spread is greater than 1% of the bid price
- The bid size or the ask size is less than zero
- The recorded trade price is less than zero

Finally, based on the rolling medians of the past three returns, we omit any data point associated with a return more than 1% away from the median centered on that minute bar. This effectively removes any data points that cause very large moves in one direction followed by an immediate large reversal. This filter generally removes fewer than five data points per year per asset.

A.1.3. Volatility Signature Plots

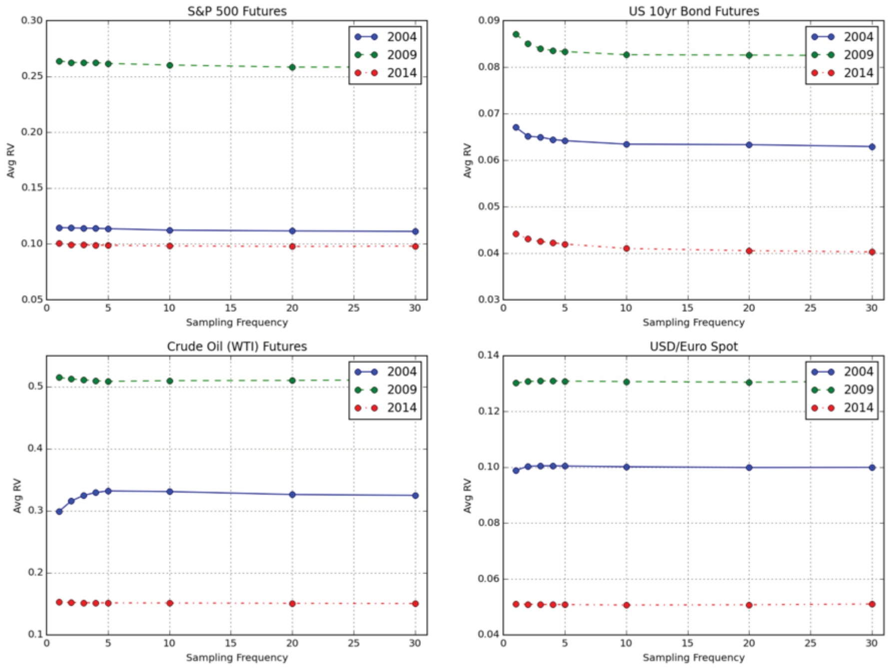
Our choice of intraday sampling frequency Δ for the high-frequency returns used in the construction of the realized volatilities is guided by the volatility signature plots first proposed by Andersen, Bollerslev, Diebold, and Labys (2000). Intuitively, in the absence of any "contaminating" market microstructure influences, the time series sample mean of the daily RV_t's provides an unbiased estimate for the true daily unconditional variance for the active part of the trading day. If, on the other hand, the RV_t's are calculated from too high a sampling frequency $1/\Delta$ to render the basic assumption of an arbitrage-free price process a good description of the discretely sampled intraday price process, the time series sample mean of RV_t will not provide an unbiased estimate of the true daily variance. Hence, by plotting the sample mean of RV_t as a function of the sampling frequency of the underlying high-frequency returns, it is possible to determine an appropriate choice of Δ. This choice could obviously differ across assets and time. However, guided by the inflection points visible in most of the signature plots for each of the individual assets, as directly illustrated by the plots for the four representative assets and three years given in Figure A.2, 5-minute naturally emerge as an appropriate common choice.

Figure A.1: **Liquidity Plots**



Note: This figure shows the proportion of business days in the relevant year (2004, 2009, or 2014) that have at least one trade in any given minute in the 24 hour day (GMT). For instance, a value of 1.0 shows for that particular minute in the day, all business days in that year had at least one valid trade, while a 0.5 would mean that only half of the days had at least one valid trade in that minute.

Figure A.2: **Signature Plots**



Note: This figure shows the average daily realized volatility for a given year (2004, 2009, or 2014) as the sampling frequency varies.