A Novel Market Sentiment Measure: Assessing the link between VIX and the Global Consciousness Projects Data

Ulf Holmberg (ulf.e.holmberg@me.com)
Independent researcher https://orcid.org/0000-0001-5872-0076

Research Article

Keywords: Stock market returns, VIX, Global Consciousness Project

Posted Date: November 16th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3614417/v1

A Novel Market Sentiment Measure

Assessing the link between VIX and the Global Consciousness Projects Data

Ulf Holmberg*

Abstract

The Standard & Poor's 500 Volatility Index (VIX), a common measure of market sentiment, is found to be significantly correlated with the Global Consciousness Projects (GCP) data. More specifically, the largest daily composite GCP data value (*Max[Z]*) is found to significantly covary with changes in VIX. The results indicate that the GCP data can help in understanding market sentiment and that daily market movements can be better comprehended by acknowledging variations in the GCP data. As such, the results suggest that the GCP data can be put to practical use by traders, which is investigated by fitting econometric models that either utilize or ignore the GCP data on daily S&P 500 returns. Highly significant interaction terms are found both with the VIX and with daily returns from markets traded in both Europe and Asia. Additionally, it is found that recognizing such interactions can explain about one percent of the econometric model's variance. To mitigate the possibility of overfitting and P-hacking, the models are put to a practical test in an out-of-sample simulation study lasting for a predefined period of one year. In the out-of-sample simulation, an artificial trader uses S&P 500 tracking instruments and trades in accordance with the econometric model's one day ahead forecasts. The results from the out-of-sample simulations suggest that GCP data can enhance daily forecasts, making it a valuable resource for traders.

Key words: Stock market returns, VIX, Global Consciousness Project

*Independent Researcher. No conflict of interest. Email: ulf.e.holmberg@me.com.

1. Brief introduction

Prior studies have underscored a noteworthy correlation between the aggregated Global Consciousness Project (GCP) data metric, *Max[Z]*, and the daily stock market returns. However, the fundamental reasons behind the found correlation have not yet been addressed. This research endeavours to propose a hypothesis suggesting that this connection might be rooted in market sentiment, potentially influenced by events that could be picked up by variations in the GCP data. To examine this hypothesis, an analysis is conducted to investigate the covariation between the VIX, a widely utilized measure of market sentiment, and the daily GCP data metric *Max[Z]*. Furthermore, the study aims to validate the results through an out-of-sample simulation study and delves into the practical implications that such a correlation could provide for traders.

The structure of this paper is organized as follows: The subsequent section provides an in-depth discussion of both the VIX measure and the GCP data, including an exploration of their correlation. This is followed by a section in which a connection between the VIX, the GCP data, and daily stock market returns is established. Following this, an out-of-sample simulation study is presented, validating the findings by illustrating potential real-world applications of the findings by traders. To conclude, the final section summarizes discusses the results and its implications.

2.The VIX measure and the GCP data

Daily stock market returns have demonstrated a correlation with a metric derived from the GCP data (Holmberg; 2020; 2021). The driving forces behind this covariation, however, have been left for future research to explore, although it has been suggested that the dependence of daily returns on market sentiment could be a contributing factor. This, as market sentiment, in turn, may be influenced by events that the GCP data "picks up". Therefore, market sentiment could be thought of as a potential link between the GCP data and stock market returns.[1]

Market sentiment aims to describe investors' general attitudes and moods towards financial markets. This mood has been shown to correlate with trading volumes and market returns—an association that noise traders tend to acknowledge (So and Lei; 2015). Positive and negative market sentiments could thus drive price movements, even though precisely quantifying market sentiment remains challenging due to its elusive nature. Nevertheless, market participants often consider the CBOE Volatility Index (VIX), a measure of the implied volatility of 30-day S&P500 options, as a proxy (Edwards & Preston, 2017). The VIX is also often referred to as the "fear index" due to its historical correlation with market panics. VIX is calculated using the two nearest expiration months of S&P 500 options in order to achieve a rolling 30-calendar-day period.[2]

The GCP is an international and multidisciplinary collaboration project that generates and collects random number data continuously from a network of physical Random Number Generators (RNG:s). The random numbers are generated using physical processes such as avalanching and quantum tunnelling and the hypothesis underpinning the GCP suggests that events triggering widespread emotions or capturing simultaneous attention from large numbers of people may significantly influence the output of the hardware-generated random numbers in a statistically significant way.

[1] For a more detailed discussion on events impact on market sentiment, refer to e.g., Jordà's (2005) and Fraiberger et al.'s (2018).

[2] A more detailed description on the VIX index can be found in the Appendix.

The idea that the mind can affect matter, such as RNG:s, has a long history in science and studies conducted by the GCP have also yielded results that validate the project's hypothesis (see, for instance, Nelson et. al.; 2002, Radin; 2002 and Nelson and Bancel; 2011 among others). However, because the possibility that the mind can affect the random numbers produced by the GCP seems to challenge the current understanding of physics, the results have been criticised (Scargle; 2002) and most scientists demand a high standard of evidence.

It has been suggested that the GCP results are due to the experimenter selecting events supportive of the project's hypothesis and May and Spottiswoode (2011) suggested that the source of the statistical deviations reported could be attributed to a psi-mediated experimenter effect. Bancel (2011) however, analysed the data thoroughly and rejected the simple selection hypothesis with a reasonably high level of confidence. Furthermore, studies made on the correlation between the GCP data aggregate *Max[Z]* and daily stock market returns (Holmberg; 2020 and 2021), as well its relationship with global internet search trends (Holmberg; 2023) have produced results supportive of the hypothesis underlying GCP.

But what is *Max[Z]* and how is it calculated? Let the data produced by an individual physical random number generator (*RNG)* be denoted $RNG_{i,\tau}$ for $i = 1,2, \dots n_\tau$, where n_τ is the total number of operating *RNGs* during second $\tau \in t$. The *RNGs* used by the GCP produce a series of 200 bits per second with an expected value of $\mu = 100$ and a variance of $\sigma^2 = 50$ from which n_τ standardized random numbers ($z_{i,\tau}$) can be calculated.

The intraday data is aggregated into daily data bundling the GCP data into 15-minute (900 seconds) nonnegative data chunks and by applying the following formulas to the extracted data:

$$Z_\tau = \left| \sum_{\tau-900}^{\tau} z_\tau \Big/ \sqrt{900} \right|, \tag{1}$$

with

$$z_\tau = \left(\sum_i^{N_\tau} RNG_{i,\tau} - N_\tau \mu \right) \Big/ \sqrt{N_\tau \sigma^2} \tag{2}$$

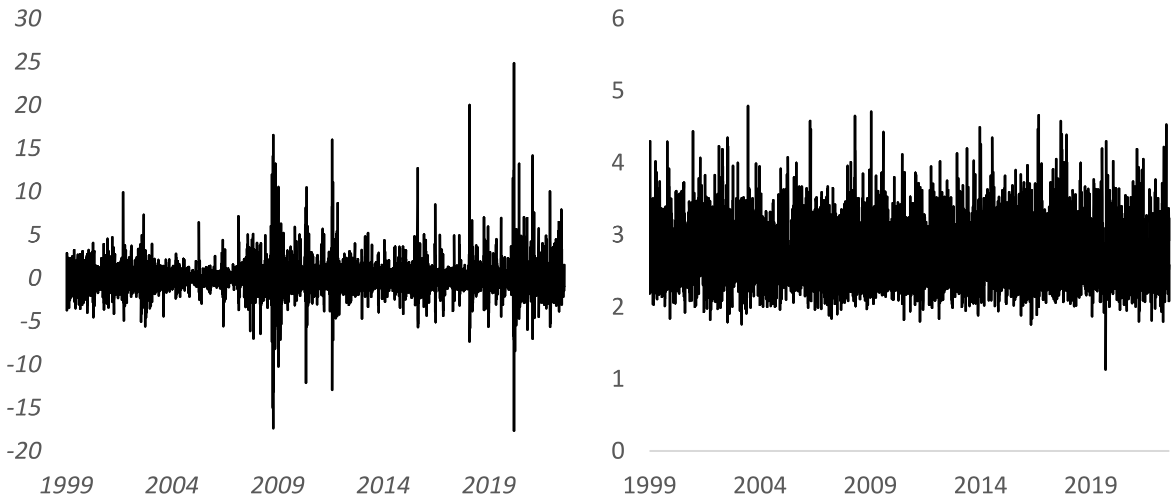
and were N_τ is the number of active RNGs during τ. Measuring Z_τ at the end of each 15-minute interval, as is done in the daily tables section on the GCP website, 96 daily intraday measurements are obtained, and from these measurements, a daily maximum is calculated. The daily maximum value of Z_τ should conceptually capture large and unexpected values in Z_τ, which in turn could covary with market sentiment measures, such as the VIX index. [3]

Historical data on VIX between 1999-01-04 and 2022-08-01 is aligned with the daily GCP data aggregate *Max[Z]*.[4] Table 1 presents descriptive data on both the *ΔVIX* and *Max[Z]* and Figure 1 depicts the two variables. Observing the data, it becomes evident that there are two notably significant "spikes" in VIX. The first spike occurred during the major financial crisis of 2008, and the second spike followed the Covid-19 pandemic in 2020. Noteworthy spikes also align with other historical events, including the Asian financial crisis in late 1997, the Russian and LTCM crisis in late 1998, the 9/11 terrorist attacks, and the European sovereign debt crisis of 2011-2012. Furthermore, upon visual examination there are

[3] Arguably, other intraday time frames could have also been chosen. However, large measurable intraday movements caused by engaging global events should "show up" in the aggregation procedure regardless of the exact time frame chosen.
[4] Dates with *Max[Z]* values larger than 5 are removed, as well dates with malfunctioning *RNG*: s.

indications that ΔVIX might exhibit heteroskedasticity, a hypothesis later confirmed through a statistical analysis, and additionally, ΔVIX demonstrates autocorrelation. *Max[Z]* also displays certain prominent "spikes," although their prominence is diminished due to its inherently greater noise. Both variables however showcase random fluctuations around their respective means. Consequently, the econometric analysis is deliberately structured to explore whether these variables simultaneously revert back to their mean values and if the mean reversion process is affected by their covariation.

Figure 1: Changes in the VIX index and *Max[Z]*



a. ΔVIX *b. Max[Z]*

Note: Daily data (N=5749) collected between 1999-01-04 and 2022-08-01.
Source: Macrobond, the Global Consciousness Project and own estimates.

Table 1: Descriptive statistics, *ΔVIX* and *Max[Z]*

	ΔVIX	***Max[Z]***
Average	-0.00	2.75
Median	-0.08	2.70
Std. Dev.	1.84	0.41
Minimum	-17.64	1.13
Maximum	24.86	4.79
Skewness	1.48	0.73
Kurtosis	24.27	1.19

Note: *Daily data (N=5749) collected between 1999-01-04 and 2022-08-01.*
Source: Macrobond, the Global Consciousness Project and own estimates.

The hypothesis postulates that *Max[Z]* might encapsulate an undisclosed facet of market sentiment, thus potentially leading to their correlation. Considering the autocorrelation of ΔVIX, the investigation into this correlation is approached using the subsequent linear equation:

$$\Delta VIX_t = \alpha + \beta_1 \Delta VIX_{t-1} + \beta_2 I_t + \sum_i \gamma_i Max[Z_{t-i}] + \sum_j \delta_i (\Delta VIX_{t-1} Max[Z_{t-i}]), \tag{3}$$

where I_t is an indicator variable equal to unity on Mondays or if the previous days data has been removed.[5] The significant correlations are tested for using the t-statistic on the models' parameters obtained through Ordinary Least Squares (OLS).

Table 2 provides the OLS regression estimates, revealing significant covariance between VIX and several interaction terms (δ_i), thus linking VIX with *Max[Z]*. Specifically, the change in ΔVIX from the previous day demonstrates significant correlation with the current day's *Max[Z]* as well as its lags for up to three days (*P < 0.01*). This indicates that if an event is detected by the GCP data, resulting in an elevated *Max[Z]* value, market sentiment is influenced for multiple consecutive days.

The sign of the parameters unveils insights into the dynamics: when market sentiment is on the upswing (indicated by a negative ΔVIX_{t-1})), a significant event resulting in a high *Max[Z]* value reverses this trend. This holds true also if the GCP data affecting event occurred on the very day that market sentiment was improving (day t-1). However, if the event occurred before market sentiment had begun to improve (i.e., during t-2 or earlier), the sign is positive indicating that the improvement instead was accelerated. It thus seems like the event that got picked up by the GCP data affects VIX differently if market sentiment is improving or deteriorating. Note that the outcomes in Table 2 thus harmoniously align with the findings from Holmberg (2021) in which it was shown that *Max[Z]* contributes positively to today's returns, only when yesterday's returns were negative and vice versa.

Table 2: ΔVIX_t model estimates

Variable / Model	Control	1	2	3	4	5	6
Constant	-0.03	0.00	0.21	-0.04	-0.03	-0.26	-0.02
$\boldsymbol{\Delta VIX_{t-1}}$	-0.15***	0.26***	0.07	-0.46***	-0.49***	-0.15	-0.24***
$\boldsymbol{Max[Z_t]}$	-	-0.01	-				-0.01
$\boldsymbol{Max[Z_{t-1}]}$	-	-	-0.09				-0.09
$\boldsymbol{Max[Z_{t-2}]}$	-	-	-	0.00			0.00
$\boldsymbol{Max[Z_{t-3}]}$	-	-	-		0.00		0.01
$\boldsymbol{Max[Z_{t-4}]}$	-	-	-			0.08	0.09
$\boldsymbol{\Delta VIX_{t-1} \times Max[Z_t]}$	-	-0.15***	-				-0.16***
$\boldsymbol{\Delta VIX_{t-1} \times Max[Z_{t-1}]}$	-	-	-0.08**				-0.08***
$\boldsymbol{\Delta VIX_{t-1} \times Max[Z_{t-2}]}$	-	-	-	0.12***			0.11***
$\boldsymbol{\Delta VIX_{t-1} \times Max[Z_{t-3}]}$	-	-	-		0.13***		0.13***
$\boldsymbol{\Delta VIX_{t-1} \times Max[Z_{t-4}]}$	-	-	-			0.00	0.03
$\boldsymbol{I_t}$	0.54***	0.54***	0.54***	0.54***	0.54***	0.54***	0.53***
$\boldsymbol{R^2}$	2.70%	3.05%	2.83%	2.93%	2.95%	2.74%	3.74%

*Significance levels: * 10%, ** 5% and *** 1%.*
Note: The daily data (N=5749) was collected on data between 1999-01-04 and 2022-08-01.

Given that investor sentiment notably influences stock markets (as explored in works like Brown and Cliff, 2005), the outcomes detailed in Table 2 offer insight into the prior discovery—that daily stock market returns is correlated with *Max[Z]*. Consequently, it appears plausible that *Max[Z]* could be

[5] Some observations are removed due to technical malfunctions distributing in the underlying RNG data used for the *Max[Z]* calculations.

capturing certain market-affecting information not presently accounted for by VIX. This suggests that the GCP data could be put to practical use my market participants.

3. VIX, *Max[Z]* and daily stock market returns

The practical implications of the findings in Table 2 are explored by estimating two econometric models on the daily S&P 500 return (r_t). The first model disregards the GCP data ($r_{t,\,without}$), while the second model incorporates its influence ($r_{t,\,with}$). Subsequently, the performance of a GCP data-dependent model is contrasted with the performance of an almost identical GCP data-independent counterpart. Notably, both models are adjusted to account for known influential factors.

A linear time series regression model is specified that allows for autocorrelated returns:[6] Also, following the insights from Pagan and Schwert (1990), Rogers et al. (1994), and Ghysels et al. (2006), the volatility in returns, a crucial factor in comprehending daily stock market dynamics, is represented by squared lagged S&P 500 returns. To accommodate the observation that investor sentiment affects stock markets, the models are designed to allow returns to be correlated with yesterday's ΔVIX. Recognizing that market sentiment tends to have a global impact, cross-market correlations in returns are also considered.

It is assumed that the Asian market is influenced by the previous day's US return, while European markets are influenced by both the previous day's and the current day's developments in the US, as well as the current day's performance in Asia. Consequently, the models embrace the notion that the S&P 500's behaviour is influenced by European and Asian returns, alongside market volatility and market sentiment. Thus, it is assumed that the performance of the US markets is not solely autocorrelated but intricately interlinked with the performance of other markets, in ways that are intricate and not always transparent.[7]

In practice, trading activities in Asia are factored into the analysis using two well-established indices: the Japanese Nikkei 225 index, traded at UTC+9, and the Hong Kong Hang Seng index, traded at UTC+8. The European market is considered, slightly overlapping with the US stock market, through the inclusion of the OMXS-30 index, traded at UTC+1.[8] Furthermore, the well-documented Monday effect (Cross, F. 1973) is addressed by incorporating an indicator variable set to 1 for Mondays. Additionally, the Monday effect is captured through numerous interaction terms involving the indicator variable. Finally, the variables' lagged dependencies are determined from the data.

To explore whether and how the GCP data aligns with daily stock market returns, an analogous econometric model is constructed. Within this model, the variables are permitted to interact with past *Max[Z]* values, enabling an investigation into the potential correlations between the GCP data and

[6] Serial correlated returns is likely since market wide information tends to get incorporated gradually causing serial correlation in the short term (see e.g., Chordia and Swaminathan; 2000, Sias and Starks; 1997, and Lo and MacKinlay; 1990 for a more detailed discussion).
[7] Since the New York Stock Exchange accounts for about half if the global market capitalization, daily market sentiment can be said to "reset" when markets in the US open for business at 14:30 (UTC). The change in sentiment could thus also affect market prices in Europe, and the intraday trend after US markets open.
[8] The Swedish stock market has a high degree of foreign ownership and closes at UTC 16:00 i.e., about 1.5 hours after the New York Stock Exchange opens (14:30 UTC). It is thus an ideal index as it then also captures a possible "reversal" and "reset" of daily sentiment once the US markets opens.

daily stock market movements. Both models are designed such that they can be used for one day ahead forecasts.[9]

Table 3 presents the model estimates, obtained on using OLS on data from between 1999-01-04 and 2022-08-01. The parameters significance is tested for using the t-statistic. The “Control model” is fitted without any GCP data dependence (i.e., $r_{t, without}$) and the “GCP data model” contains the hypothesised *Max[Z]* dependence (i.e., $r_{t, with}$).

The results in Table 3 are highly informative. As expected, the results distinctly demonstrate the autocorrelation of daily returns ($S\&P500_{t-1}$), their dependence on market variance ($S\&P500_{t-1}^2$), as well as their responsiveness to market sentiment (ΔVIX_{t-1}). Moreover, the returns are influenced by the performance of both the European ($Sweden_t$) and Asian markets ($Hong\ Kong_t$ and $Japan_t$) in addition to being influenced by past returns across global markets ($Sweden_{t-1}$ and $Japan_{t-1}$). Additionally, it's evident that Monday returns are different, substantiated by the high significance of various interaction terms involving the binary Monday indicator (*P < 0.01*).

Even more intriguing findings emerge when examining the correlation between the GCP data (*Max[Z]*) and S&P 500 returns. Although the significance of lagged *Max[Z]* is absent when considered on its own alongside other variables, it dynamically interacts with OMXS-30 returns (Sweden), Hang Seng returns (Hong Kong), and also responds to fluctuations in market sentiment (ΔVIX_{t-1}). Notably, all these interaction terms carry substantial significance (*P < 0.01*). While the interaction term involving Nikkei 225 returns (Japan) is initially non-significant, its significance surfaces when excluding the developments in Hong Kong and Stockholm. This result implies that the Nikkei 225 index also interacts with *Max[Z]*, albeit in a manner better captured by the other interaction terms. Moreover, the introduction of *Max[Z]* interaction terms amplify the explained variance within the model. This augmentation is highlighted by an increase in the coefficient of determination (R^2) by more than 1 percent. Importantly, this enhanced explanatory power remains evident even after accounting for the additional parameters introduced into the model (R_{adj}^2).

The insights gleaned from Table 3 suggests that daily market movements can be better understood by acknowledging information contained within the GCP data. Consequently, the results suggest a practical utility for traders in utilizing GCP data. However, it could be claimed that the results are due to data-fitting such that the correlations alone are not sufficient to substantiate any assertions. In response to such claims, an "out -of-sample" simulation study, spanning over a predefined period of one year, was conducted. During the simulation, one day ahead forecasts were generated using the parameters in Table 3 and an artificial trader was assumed to operate in alignment with the derived estimates. The simulation thus serves as a means to address if the significant GCP data interactions are valid and to study if the results can be effectively applied in real-world trading scenarios.[10]

[9] In Holmberg (2020 and 2021), also todays *Max[Z]* was found to correlate with today’s return, an intuitive finding as the GCP data reacts to events affecting daily stock markets directly. However, as the model is results are validated in an out of sample simulation using on one day ahead forecasts, only interactions with past returns are included.

[10] The out of sample simulation was onset in August 2022 and made public continuously on the authors webpage (www.ulfholmberg.info). Preliminary results from the simulation was also presented during a poster session during the TSC 2023 conference in Taormina, Italy (PO-2 (Fri): “*Consciousness, sentiment and stock market returns: Could the GCP data be put to practical use?”*.

Table 3: The S&P500 daily returns models
Standard errors in parentheses

	Without GCP data	With GCP data
$\boldsymbol{Constant}$	$1.380E-04$ $(1.540\mathrm{E}-04)$	$-1.370E-04$ (9.140E-04)
$\boldsymbol{S\&P500_{t-1}}$	-0.312^{***} (0.020)	-0.304^{***} (0.020)
$\boldsymbol{S\&P500^2_{t-1}}$	0.899^{***} (0.271)	0.882^{***} (0.270)
$\boldsymbol{Sweden_t}$	0.437^{***} (0.012)	0.546^{***} (0.068)
$\boldsymbol{Hong\ Kong_t}$	0.062^{***} (0.011)	0.604^{***} (0.078)
$\boldsymbol{Japan_t}$	0.054^{***} (0.013)	-0.071 (0.079)
$\boldsymbol{Sweden_{t-1}}$	0.073^{***} (0.013)	0.072^{***} (0.012)
$\boldsymbol{Japan_{t-1}}$	0.009 (0.011)	0.012 (0.011)
$\Delta\boldsymbol{VIX_{t-1}}$	$-1.440E-05$ $(1.280\mathrm{E}-04)$	$1.074E-03^{*}$ $(5.930E-04)$
$\boldsymbol{Monday_t}$	$-2.740E-04$ (3.450E-04)	$-3.180E-04$ $(3.430E-04)$
$\boldsymbol{Monday_t} \times \boldsymbol{Sweden_t}$	0.118^{***} (0.024)	0.123^{***} (0.023)
$\boldsymbol{Monday_t} \times \boldsymbol{Sweden_{t-1}}$	-0.071^{**} (0.028)	-0.085^{***} (0.028)
$\boldsymbol{Monday_t} \times \boldsymbol{Japan_t}$	0.050^{*} (0.027)	0.050^{*} (0.027)
$\boldsymbol{Monday_t} \times \boldsymbol{Japan_{t-1}}$	0.067^{***} (0.025)	0.060^{**} (0.025)
$\Delta\boldsymbol{VIX_{t-1}} \times \boldsymbol{Japan_{t-1}}$	-0.002 (0.003)	$2.262E-03$ $(3.499E-03)$
$\Delta\boldsymbol{VIX_{t-1}} \times \boldsymbol{Japan_{t-1}} \times \boldsymbol{Monday_t}$	-0.031^{***} (0.006)	-0.031^{***} $(6.301E-03)$
$\boldsymbol{Max}[\boldsymbol{Z_{t-1}}]$	-	1.130E-04 (3.280E-04)
$\boldsymbol{Max}[\boldsymbol{Z_{t-1}}] \times \boldsymbol{Sweden_t}$	-	-0.041^{*} (0.025)
$\boldsymbol{Max}[\boldsymbol{Z_{t-1}}] \times \boldsymbol{Hong\ Kong_t}$	-	-0.200^{***} (0.028)
$\boldsymbol{Max}[\boldsymbol{Z_{t-1}}] \times \boldsymbol{Japan_t}$	-	0.046 (0.029)
$\boldsymbol{Max}[\boldsymbol{Z_{t-1}}] \times \Delta\boldsymbol{VIX_{t-1}}$	-	$-6.940E-04^{***}$ $(2.170E-04)$
$\boldsymbol{Max}[\boldsymbol{Z_{t-1}}] \times \Delta\boldsymbol{VIX^2_{t-1}}$	-	$4.490E-06^{***}$ $(7.800E-07)$
$\boldsymbol{R^2}$	0.341	0.352
$\boldsymbol{R^2_{adj}}$	0.339	0.349

*Significance levels: * 10%, ** 5% and *** 1%.*
Note: OLS estimates on daily data (N=5749) collected between 1999-01-04 and 2022-08-01.

4. An out-of-sample simulation study

The results in Table 3 suggest that the GCP data interacts with stock market returns in ways that could be utilized by traders. However, as it in principle is possible to fit a polynomial to the data such that significant correlations are found even though no true correlation exists, this section presents results from an out-of-sample simulation study lasting for one year. The simulations where onset on the 1st of August 2022 and ended on the 31st of July 2023.[11] Also, the results from the simulations using approximate S&P 500 future prices were made public every week on the authors webpage and in order to study the effect from market pricing, the models used where disclosed after 6 months, i.e., on the 1st of February 2023.[12]

The term "out-of-sample" refers to that the models used where fitted on a different sample than the period on which the simulations were made. This as the parameters in Table 3, fitted on data collected between 1999-01-04 and 2022-08-01, were used for one day ahead S&P 500 returns forecasts during a period spanning from 2022-08-01 to 2023-07-31. By doing so, the possible problems related to "P-hacking" is addressed as if no true correlation exists, the simulations should reveal that no advantage can be made by utilizing the GCP data. However, if the fund that utilized the GCP data is shown to outperform its GCP data invariant counterpart, the GCP data's usefulness in an actual trading environment has been demonstrated.[13]

In the out-of-sample simulations, the artificial trader is assumed to either buy and hold a S&P 500 tracking instrument or to trade actively in accordance with expectations aligned with the one day ahead forecasts obtained using the parameters in Table 3. If the forecasted return is positive, the trader is assumed to buy (go long) and if the forecast is negative, the trader is assumed to sell (go short). All open positions are closed when the market closes (UTC 21:00). Table 4 summarizes the studied trading strategies.

Table 4: Investment strategies

S&P 500 (B&H)	GCP data fund	Control fund
Buy at start and hold	Long if $\hat{r}_{t,with} \geq 0$	Long if $\hat{r}_{t,without} \geq 0$
	Short if $\hat{r}_{t,with} < 0$	Short if $\hat{r}_{t,without} < 0$

Note: $\hat{r}_t$ is the one day ahead model forecast.

Note that the econometric models in Table 3 requires a close value from the OMXS-30 index. However, the market in Stockholm closes at UTC 16:00 such that trades can be made on this market for another 1.5 hours after the stock market in New York has opened for trade (UTC 14:30). As such, trades made ahead of UTC 16:00 will overestimate the simulated returns. The artificial trader could also wait until till the market closes in Stockholm (UTC 16:00), but when doing so it is assumed that the artificial trader ignores the information embedded during the day in both Europe and in Asia as. In practice, it is likely that a day trader will trade sometime between S&P 500 open (UTC 14:30) and when markets close in Sweden (UTC 16:00). As such, trades made at the open price (UTC 14.30), as well as trades made at

[11] It should be noted that it ahead of the simulations was likely that the GCP data effect would fade with time as the well known
[12] www.ulfholmberg.info
[13] It should be noted that it is highly unusual to keep the econometric models constant for a full year. However, as the funds relative performance is studied, this is of little concern.

UTC 15:00 and UTC 16:00 are investigated using the S&P 500 futures price. [14] By doing so, also the potential advantage of early action vs waiting can be studied.[15]

Note however that the out-of-sample simulation seeks to investigate if the GCP data can add value to traders by comparing GCP data dependent trades with a GCP data invariant counterpart. Since this is investigated by comparing a GCP data dependent fund with a Control Fund, the effect on the absolut fund level of trading ahead of UTC 16:00 is of little concern, as the research hypothesis can be addressed through the fund's relative performance.

On the first of August 2022, 100 currency units are made available for investments. The B&H trader immediately invests the full amount in an S&P 500 index tracking instrument and the actively traded funds make their first trade is made on the 2nd of August 2022.[16] Note also that trades are only allowed to be made on dates on which datapoints exists on variables needed for making the forecast (i.e., all listed variables in Table 3).

Figure 2 depicts the funds' performance, both in absolute terms (Figure 2.a) and in relative terms (Figure 2.b). From Figure 2.a, it can be seen that the actively traded funds in general outperform the passive B&H (S&P 500) strategy, except for the simulations in which the trades are made at UTC 16:00 between February and July 2023. Importantly, the GCP data funds value is in general higher than the Control funds, regardless of when the trade was made which is suggesting that the GCP data adds value for investors.

Focusing on the actively traded funds relative performance (i.e., the value of the GCP data fund after subtracting the value of the Control fund), it can be seen that the two funds traded ahead of UTC 16:00 quickly begun to outperform their GCP data invariant counterparts (in Figure 2.b). Even though the GCP data dependent funds underperformed during the end of October/ early November, the GCP data funds again outperformed the Control fund from November to January and during the second half of the simulation period, they mostly performed identically.

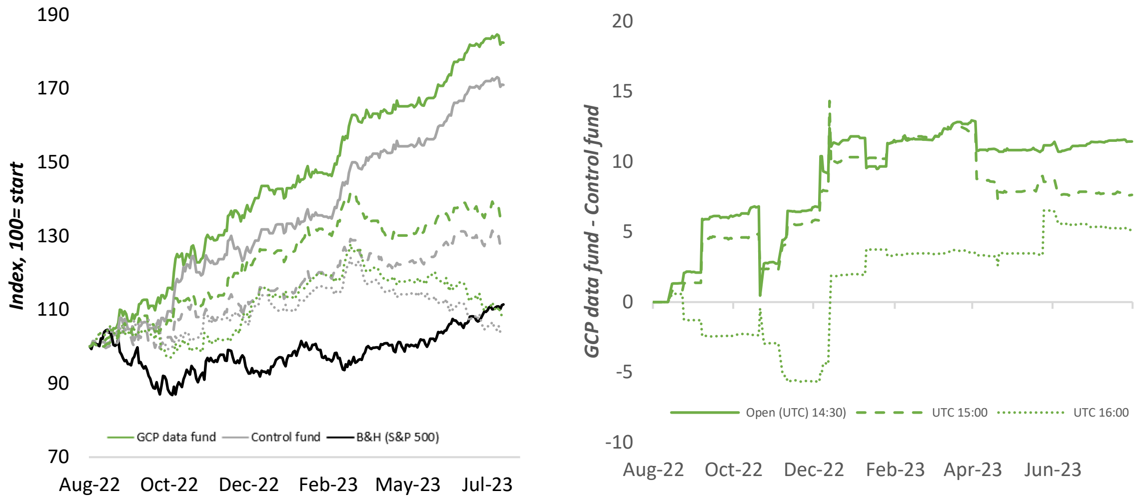
Turing to the UTC 16:00 traded funds relative performance, it can be seen that the GCP data dependent fund underperformed between September and mid-December. However, due to favorable trades made during the end of the year, also the UTC 16:00 traded GCP data fund outperforms the Control fund. When the simulations ended on the 31st of July, all three GCP data dependent funds had outperformed their control funds by several percentage points.

Table 5 presents the actively traded funds relative performance once the simulations had ended and also here, the results clearly point towards that the GCP data adds value to traders. It should however be noted that the relative return tends to decrease if the investor waits for the market in Sweden to close (UTC 16:00) and that the relative hit rate, a metric capturing the share of positive return trades, is at its highest if the trades are made at UTC 15:00. As such, the GCP data can be said to add value

[14] E-mini S&P 500 Futures prices from Firstratedata.com at EST 10:00 and EST 11:00 are used.
[15] These issues are of minor concern as the study aims to investigate if the GCP data can be used in practice which only requires a comparing with a control fund.
[16] The simulations ignore potential brokerage fees.

during market value uncertainty, a result well aligned with the finding in Holmberg (2021), i.e. that *Max[Z]* correlates to a higher degree with stock markets during periods of high market volatility.

Figure 2: Funds out-of-sample performance



a. Value of the artificial funds **b. GCP data funds relative performance (%)**

Note: Solid lines represents the open price (CET 14:30) simulations, dashed the CET 15:00 simulations and dotted the CET 16:00 simulations.

In Table 5, also a version of the simulations in which the GCP data funds investment decision is conditioned on the size of yesterday's *Max[Z]* are presented (Filtered). In these "filtered" simulations, the GCP data fund trader use the Control funds forecast if *Max[Z]* < $|N^{-1}(1\%)|$. By filtering the *Max[Z]* variable in such a way, all small and possibly irrelevant, GCP data values are ignored. As can be seen, filtering the GCP data increases the relative hit with one percent which in turn increases the relative return with between 0.8 and 4.7 percent.

Table 5: The artificial funds' relative performance

	Relative total return		***Relative hit rate****	
	Unfiltered	***Filtered***	***Unfiltered***	***Filtered***
Open (UTC 14:30)	*11.4*	*13.9*	*0.5*	*1.5*
UTC 15:00	*7.6*	*12.6*	*1.5*	*2.5*
UTC 16:00	*5.1*	*5.9*	*-0.5*	*0.5*

**Share of trades resulting in positive returns.*

Table 6 breaks down the artificial funds' relative performance into monthly contributions and here the results confirm the importance of when the trade is made. If the trade is made at UTC 16:00, the relative performance looks strikingly like chance but if trades are made slightly after the market has opened (UTC 15:00), the GCP data fund clearly outperforms the control fund. In fact, the GCP data fund produced excess returns in 6 of the 9 months, while the control fund produces excess returns in only 3. This difference is significant at the 10 percent level (*P=7.9%*) and if the data is GCP data is filtered, the P value falls to below one percent as an additional month is added to the list of month in which the GCP data fund outperformed the control fund (*P= 0.4%*).

Also, if the first half of the simulation is analyzed, it is found that the fund traded at open (UTC 14:30) performed better than the control fund and the GCP data fund traded at UTC 15:00 outperformed the control fund in 5 out of 6 months (*P≈1%*). The control fund however outperformed the GCP data fund only in October (but not for the filtered GCP data fund). However, after the 1st of February when the models were disclosed on the authors webpage, the GCP data fund either underperformed or performed in equivalence with the Control fund. [17]

Table 6: The artificial funds' monthly relative performance (%)

The filtered funds relative performance can be found in parenthesis when different

	Open (UTC 14:30)	***UTC 15:00***	***UTC 16:00***
August, 2022	2.2	1.4	-1.3
September, 2022	3.6	3.0	-1.1
October, 2022	-3.9	-2.5	-0.6
November, 2022	3.0	3.0	-2.6
December, 2022	3.6	3.7	7.8
January, 2023	-0.4	0.9	1.2
February, 2023	0.0	0.0	0.0
March, 2023	0.0	0.0	0.0
April, 2023	-1.4 (-0.7)	-3.1 (0.5)	0.0 (0.7)
May, 2023	0.1	0.4	2.6
June, 2023	-0.5	-0.9	-0.7
July, 2023	0.0	0.0	0.0
	*P values (%) from the proportions test**		
Full simulation	31.9	7.9 (0.4)	90.0 (81.4)
From August 2022 to January 2023	12.4	1.0	87.6

** A standard proportions test on the number of months the funds performed differently. Significance levels from the Gaussian distribution.*

Figure 3 depicts the relative hit rate (i.e., the Control funds hit rate subtracted from the GCP data funds hit rate). As can be seen, all GCP data funds outperformed the Control funds in August but in September, the GCP data fund traded at CET 16:00 begun to underperform. Both funds traded ahead of markets closing in Stockholm, however, continued to outperform the Control funds.

As time progressed, the GCP data funds traded ahead of UTC 16:00 relative performance begun to decline. Multiple factors need to be considered when one seeks to understand the declining effect in Figure 3 and below, some notable explanations are listed:

(i) **The results are due to chance.** This explanation seems unlikely as the statistical proportion test on the number of months the GCP data fund outperformed the Control fund suggests otherwise. The proportions tests points towards a significance at the 10 percent level if the GCP data is unfiltered and at the 1 percent level if the GCP data is filtered, which is confirmed with a proportions test using the daily data. Furthermore, the GCP data model correctly predicted more trades correctly than the control model and the final outcome of the simulations points towards the GCP data adding value to traders (Table 5).

[17] If the GCP data fund is filtered and if trades made during US Federal reserve decision weeks are ignored, the difference in proportions becomes significant at the 1 percent level over the year, and at the 10 percent level between February and July 2023.

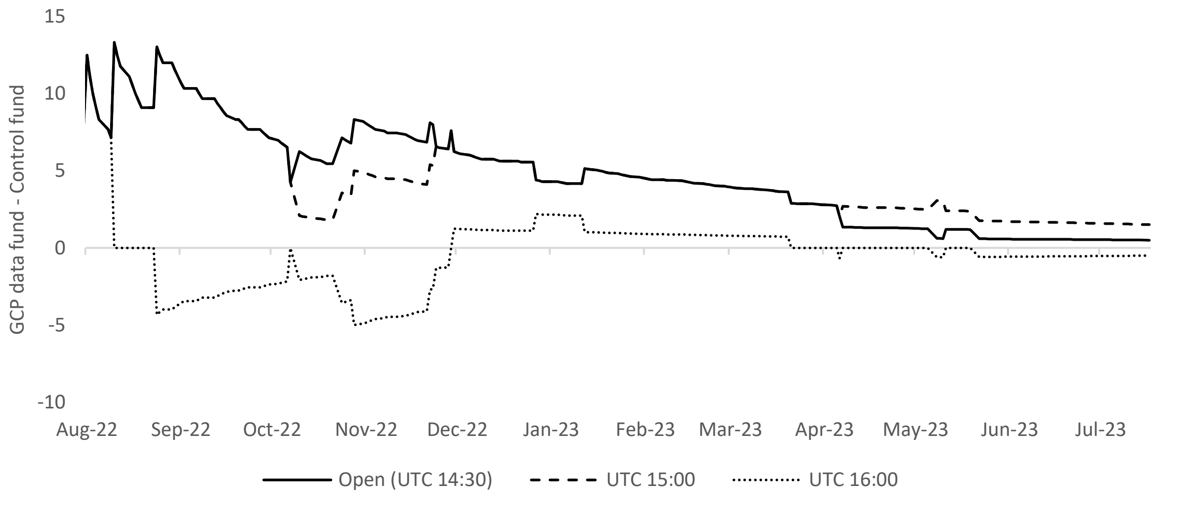
(ii) **Improved market sentiment.** Previous research has found that *Max[Z]* correlates more strongly with daily stock market returns during volatile time periods (Holmberg; 2021). As market volatility is related to market sentiment, the declining effect could the result of that VIX begun to fall during the simulation period. On average, VIX fell from 24.1 during the first half of the simulations to 17.5 during the second half. [18] The decrease in VIX signals a period of more stable returns and comparing the standard deviation in returns between the first and second half of the simulation period, the standard deviation in returns fell from 1.41 percent to 0.82 percent.

(iii) **No arbitrage**. The no arbitrage assumption suggests that the markets do not allow for risk-free profits with no initial investment. This as any difference in returns that can be obtained without taking on risk, will be traded away as prices adjust. Since the models were made public on the 1st of February, future prices could have been affected which in turn could have influenced the results.

(iv) **The declining Psi effect.** The *Max[Z]* variable is calculated out of the GCP data, and several research papers have found that the Psi effect tends to decline with time (see e.g., Bierman et.al., 2016 and Radin, D. 2006 among others).

(v) **Complex links.** The link between the GCP data, *Max[Z]* and stock market returns could be more complex than the other links acknowledged. As such, the GCP data dependent model could be in need of more frequently updates which in turn could have influenced the results.

Figure 3: Relative hit rate (%)



Returning to Figure 4, VIX alone is able to explain between 43 and 58 percent of the observed decline during the year.[19] However, the complex ways in which the GCP data interacts with markets as well as the fact that the models were made public after 6 months can both be affecting the results. The exact

[18] The decrease in VIX coincided with the Zero-Dated Options debate in which it was argued that zero-dataed options had "broken" VIX. However, as VIX reflects the market's best estimate of SPX volatility over the coming 30 days, this seems unlikely (Sosnicks, 2023).

[19] Simple linear regressions on relative performance of the Open (CET 14:40) and CET 15:00 traded funds suggests that 58 and 43 percent of the model variance can be explained by VIX alone.

reason for the decline is thus likely to be complex and left as an interesting avenue for future research to explore.

5. Concluding remarks

In this paper, the GCP data's covariation with market sentiment, and how this correlation can be but to practical use by market participants, has been explored. The paper begun with correlating the *Max[Z]* variable, a daily aggregate derived out of the GCP data, with the commonly used VIX measure. It was found that not only did todays *Max[Z]* correlate with changes in VIX (*P< 0.01*), but so did also several of its lags. As investor sentiment is known to have a measurable impact on stock markets (see e.g., Brown and Cliff; 2005), the results shed light on the unorthodox finding in Holmberg (2020 and 2021), i.e., that that the GCP data covaries significantly with daily stock market returns.

This is a striking result. Not only does it point towards the validity of the hypothesis underlying the GCP data, but it also suggests that the data can be used in practice by e.g., traders. This is studied by fitting two almost equivalent econometric models on daily S&P 500 returns, one that uses the GCP data and one that does not, where the latter is used as the control model. Both models are designed in ways that allow for one day ahead forecasting. Here, it is again found that *Max[Z]* covaries with daily returns as yesterday's *Max[Z]* value significantly covaries with today's OMXS-30 returns, with Hang Seng (Hong Kong) returns, as well with changes in yesterday's VIX and with its variation (ΔVIX_{t-1}^{2}). Furthermore, it is found that about one percent of the econometric models' variance is explained by these interaction terms which suggests that the results can be used to gain a competitive edge in markets.

The potential advantage of using the GCP data is studied in an out-of-sample simulation. The simulations were set to last of one year and onset on the 1st of August 2022. Trades made during three different time periods were studied and when the simulations ended on the 31st of July 2023, the results clearly showed that the GCP data can be used to inform traders. In fact, if the trades were made when the market opened in New York (UTC 14:30), the GCP data informed trader achieved between 13.9 and 11.4 percent higher annual returns then its GCP data invariant counterpart. Furthermore, if the artificial trader waited for half an hour and traded at UTC 15:00, the GCP data had the ability to increase annual returns with 12.6 and 7.6 percent.

It could however be argued that the out-of-sample simulations in which the investment decisions were made at UTC 14:30 or UTC 15:00 overstate the GCP data's contribution to the fund's total return. This as OMXS-30 closes at UTC 16:00 such that the artificial trader needs to forecast the OMXS-30 indexes movements during the final hours of trade. However, as both the GCP data fund and the control fund are subject to this shortcoming, this issue can be ignored such that the results are solid and point towards the usability of the GCP data. For completeness however, also simulations from trades made after the market closes in Stockholm are studied and also then the, GCP data fund is found to outperform the Control fund.[20]

Taken together, the out-of-sample simulations suggest that investors can gain an edge by acknowledging the information embedded in the GCP data. However, it was also found that the edge gained decreases with the time the trader waits with making the daily trade as the potential gain gets

[20] The relative hit rate increases with 0.5 percent such that the relative annual return ends up being between 5.9 (using filtered *Max[Z]*) and 5.1 (using unfiltered *Max[Z]*) higher.

priced away. However, considering the results together with the findings in Holmberg (2023), i.e., that the GCP data covaries with internet search trends, the results highlight how useful the GCP data can be to forecasters in general.

It should be noted that the econometric models used in the out-of-sample simulations utilize how yesterday's *Max[Z]* covaries with markets, not today's *Max[Z]*. As the results from the analysis presented in here also shows a correlation between yesterday VIX and todays *Max[Z]* and since previous research has found a correlation between daily returns and todays *Max[Z],* it is possible that even more accurate and timely econometric models can be constructed, models that account for also the current days intraday *Max[Z]* movements. Furthermore, it is possible that more accurate daily GCP data effect measures can be constructed, measures that covary even more strongly with markets. How such models or measures can be constructed is however left as an interesting avenue for future research to explore.

As can be understood from the above, the results herein point in the direction of multiple other avenues for future research. Firstly, they suggest that the data produced by the GCP can be put to practical use by forecasters. Also, as they validate several claims made by the GCP, they also pave the way for research on alternative theories on the nature of consciousness. Furthermore, as stock market returns, sentiment, and focused attention tends to be tightly linked to the economic performance in general, it is also likely other economic metrics could be better understood by acknowledging the information embedded in the GCP data.

Finally, as with any study, limitations should be acknowledged. It is not claimed that a causal relationship between GCP data and stock market returns is established, as no theory yet exists that can explain such a link. Nonetheless, the findings suggest that GCP data could be a valuable tool for traders and provide useful insights into market sentiment dynamics. As the GCP data rests on the hypothesis that our emotions and thoughts can affect matter at a distance, both the philosophical and practical implications of these results are self evident.

References

Bierman, D. (2000). On the nature of anomalous phenomena: Another reality between the world of subjective and the objective world of physics? In The Physical Nature of , ed. by P. van Loocke. New York: Benjamin Publication.

Brown, Gregory W., and Michael T. Cliff. (2005). "Investor Sentiment and Asset Valuation." The Journal of Business, 78(2), pp: 405–40.

Carroll, R.T. (2003). The Skeptic's Dictionary: A collection of strange beliefs, amusing deceptions and dangerous delusions. London: Wiley.

Cross, F. (1973). "The behavior of stock prices on Fridays and Mondays". Financial Analyst Journal: 4: 67-69.

Chordia, T. and Swaminathan, B. (2000), Trading Volume and Cross-Autocorrelations in Stock Returns. The Journal of Finance, 55: 913-935.

Edwards, T. and Preston, H. (2017). "A Practitioner's Guide to Reading VIX?", S&P global, strategy 201

Fraiberger, S. P., Do Lee, Damien Puy and, Romain Ranciere. (2018). "Media Sentiment and International Asset Prices". International Monetary Fund Working, WP/18/274.

Güzeldere, G. (1997). The Nature of : Philosophical debates. Cambridge, MA: MIT Press: 1–67.

Ghysels, E., Sinko, A. and Valkanov, R. (2006). MIDAS regression: Further results and new directions. Econometric Reviews, 26, pp.53-90.

Jordà, O. (2005). "Estimation and Inference of Impulse Responses by Local Projections, American Economic Review 95(1), pp.161-182

Lo, A. W. and MacKinlay, A. C. (1990). "When Are Contrarian Profits Due to Stock Market Overreaction?".The Review of Financial Studies, 3 (2), pp. 175–205

Nelson, R. and Bancel, P. (2011). "Effects of Mass in Random Data During Global Events". Explore; 7, pp. 373-383.

Nelson R., Radin D., Shoup R, Bancel P. (2002). " Correlations of continuous random data with major world events". Found Phys Lett. 15, pp: 537–550.

Pagan, A.R., and Schwert, G.W. (1990). "Alternative models for conditional models for conditional stock volatility", Journal of Econometrics, 45 (1-2), pp. 267-290.

Radin, Dean (2002). "Exploring relationships Beween Random Physical Events and Mass Human Attention: Asking for Whom the Bell Tolls". Journal of Scientific Exploration. 16 (4): 533-547.

Radin, D. (2006). Entangled Minds: Extrasensory Experiences in Quantum Reality. London: Paraview Pocket Books.

Rogers, L., Satchell, S. and Yoon, Y. (1994). Estimating the Volatility of Stock Prices: A Comparison of Methods That Use High and Low Prices, Applied Financial Economics, 4(3), pp. 241-47

Scargle, J. D. (2002). "Was there evidence of global on September 11, 2001?". Journal of Scientific Exploration. 16 (4): pp. 571–577.

Sias, R. W and Starks, L. T. (1997). "Return autocorrelation and institutional investors", Journal of Financial Economics, 46 (1), pp. 103-131

So, S.M. and Lei, V. (2015). On the relationship between investor sentiment, VIX and trading volume. Risk governance & control: Financial markets & institutions, 5(4-1), 114-122

Sosnick, S. (2023). "Have Zero-Dated Options Broken VIX?", Posted on Traders Insight, 13 April 2023

Holmberg, U. (2020). Stock Returns and the Mind: An Unlikely Result that Could Change Our Understanding of , "*Journal of Studies*", 27 (7-8), pp. 31-49.

Holmberg, U. (2021). Revisiting Stock Returns and the Mind: Digging Deeper into the Data, "*Journal of Exploration & Research*", 12(3), pp. 207-223.

Holmberg, U. (2023). "Validating the GCP data hypothesis using internet search data", *EXPLORE*, 19(2), pp. 228-237,

Appendix: Calculating VIX

The VIX is computed by considering the two closest expiration months of S&P 500 options, ensuring a continuous rolling 30-calendar-day timeframe. The index value is determined based on the prices of both at-the-money and out-of-the-money calls and puts.[21] The calculation assigns greater weight to the prices of options with strike prices closer to the at-the-money value. The formula for calculating the VIX is as follows:

$$\sigma^2 = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i) - \frac{1}{T}\left[\frac{F}{K_0} - 1\right]^2, \quad (A.1)$$

where σ is the value of the VIX divided by 100, T is the time to expiration of the option contract, F is the forward index level derived from option prices, K_i is the strike price of the i^{th} out-of-the-money option (call if $K_t > F$ and put if $K_t < F$), ΔK_i is the interval between strike prices, or $(K_{i+1} - K_{i-1})/2$, K_0 is the first strike below F, r is the risk-free interest rate up to the expiration of the option contract, and $Q(K_i)$is the midpoint of the bid-ask spread for each option with strike K_i. Calls and puts are included up to the point where there are two consecutive strike prices with a bid price equal to zero. For the study, daily values for VIX are downloaded from Macrobond.

[21] At-the-money and out-of-the-money calls and puts are used to explain the positioning and rationale behind the pricing of an option contract. Near term (23 day) and Next term (>30 day) at-the-money and out-of-the-money options are weighed by time to maturity and moneyness to interpolate a 30-day price of volatility. The rollover to the next expiration occurs eight calendar days prior to the expiry of the nearby option.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Datapaper.xlsx