Journal of Banking and Finance xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Biased risk perceptions: Evidence from the laboratory and financial markets ☆,☆☆

Elise Payzan-LeNestour[a,*], Lionnel Pradier[a], Tālis J. Putniņš[b,c]

[a] UNSW Business School, University of New South Wales, NSW 2052, Australia
[b] UTS Business School, University of Technology Sydney, PO Box 123 Broadway, NSW 2007, Australia
[c] Stockholm School of Economics in Riga, Latvia

ARTICLE INFO

Article history:
Received 14 November 2021
Accepted 16 September 2022
Available online xxx

JEL classification:
D91
G02
G41

Keywords:
Behavioural finance
Neurofinance
Risk perception
Decision making under uncertainty
Efficient coding

ABSTRACT

Applying a well-established neuroscience framework to the issue of investor perception of volatility, we propose that after prolonged exposure to high volatility, investors tend to underestimate volatility due to adaptation to the high volatility, and vice versa. Using a combination of field and laboratory tests, we find strong support for this hypothesis. The evidence suggests that this neurobiologically-grounded perceptual bias can cause distortions of asset prices in sophisticated and liquid financial markets.

1. Introduction

Risk plays a fundamental role in almost all areas of finance, and a growing strand of finance research aims to understand how investors perceive risk and how well this perception lines up with objective measures of risk.[1] Following the lead of these studies, researchers have recently examined investor perception of risk through the lens of "*efficient coding theory*", the neuroscientific framework commonly used to study how the brain records information about the external world.[2] The basic idea is that the brain is designed to communicate information in a way that economizes on its limited resources (the finite number of neurons, each with a finite number of spike outputs). While efficient coding is optimal overall, it leads to some predictable errors in perception. [3]

In this paper we focus on the perceptual errors caused by an aspect of efficient coding known as "adaptive normalization": the perception of a stimulus depends not on its absolute intensity, but rather on its intensity relative to the recent past and neighbouring stimuli. As a result, perception is sensitive to contrasts or deviations from the mean (Glimcher, 2014), giving rise to "*aftereffects*", in which prolonged exposure to a given stimulus level sys-

☆ We thank Yacine Ait-Sahalia, Sean Anthonisz, Mike Aitken, Ivars Austers, Giovanni Barone-Adesi, Tony Berrada, Gilles Chemla, Claudia Custodio, Mark Dean, Rajna Gibson, Chris Hansman, Samuel Hartzmark, Terry Hendershott, David Hirshleifer,, Marcin Kacperczyk, Petko Kalev, Egon Kalotay, Semyon Malamud, David Miles, Nathalie Moyen, Phong Ngo, Olivier Scaillet, Andrei Shleifer, Shuo Song, Savi Sundaresan, Katrin Tinn, Michael Woodford, and seminar participants at the London School of Economics, Imperial College, Geneva University, the Cognition and Decision Lab at Columbia University, the Q-Group Colloquium, Stockholm School of Economics, and the 6th Vitznau Neurofinance Conference for helpful comments and insights. On the neuroscience side, we thank Giorgios Christopoulos, Paul Glimcher, Scott Huettel, John O'Doherty, and Wolfram Schultz, for extremely useful insights. We are grateful both to the Securities Industry Research Centre of Asia-Pacific and Thomson Reuters for providing access to the financial data used in this study, and to the University of New South Wales for funding our laboratory experiment.

☆☆ The Internet Appendix that accompanies this paper can be found at https://tinyurl.com/yafapd4f.

* Corresponding author.
E-mail address: elise@unsw.edu.au (E. Payzan-LeNestour).

[1] For example, Weber et al. (2013), Huber et al. (2019), and Huber et al. (2021), among others (more in "Related Literature").

[2] For example, Barlow (1972), Laughlin (1981), and Glimcher (2014).

[3] For example, efficient coding leads the brain to perceive tail events as less extreme than they actually are, possibly resulting in neglect of financial risk (Payzan-LeNestour and Woodford, 2022).

https://doi.org/10.1016/j.jbankfin.2022.106685
0378-4266/

tematically creates the illusion of an opposite stimulus. For example, after prolonged viewing of the downward flow of a waterfall, static rocks to the side appear to ooze upward (Barlow and Hill, 1963). After-effects have been observed for the perception of many other stimulus properties, from colour (Hurvich and Jameson, 1957) to abstract properties such as the perceived numerosity of dots in patches (Burr and Ross, 2008), the perception of economic value (Khaw et al., 2017), and the perception of corporate earnings (Hartzmark and Shue, 2018). Applied to investor perception of volatility, the basic idea is that after prolonged exposure to very high volatility, investors may perceive neutral ("medium") volatility levels as lower than actual, and vice versa—the perception of medium volatility may be biased upward after prolonged exposure to very low volatility levels.

To test this idea, Payzan-LeNestour et al. (2016) asked participants in a laboratory experiment to rate the volatility of a medium volatility (0.1) asset displayed on screen after prolonged exposure to very high volatility (VH = 0.45) in half the trials, and very low volatility (VL = 0.02) in the other half. Participant ratings were on average lower after exposure to VH ("post-high") than after exposure to VL ("post-low"), consistent with the idea that their perception of volatility was biased by the after-effect.

In this study, we turn to a combination of laboratory experiments and field data to study the pervasiveness and characteristics of the volatility after-effect. We test for asset price distortions caused by the after-effect. We reason that the existence of such distortions would show that the after-effect affects the volatility perception of the marginal trader, beyond the average individual in the laboratory. The methodology used in the study is best described as a "loop process" whereby we start in the lab, then turn to the field to check the external validity of the lab results, and ultimately return to the lab to better understand features of the field evidence. This approach illustrates the complementarity of laboratory and field data (List, 2007).

Our empirical tests use the Market Volatility Index (VIX), which embeds investor forecasts of future volatility. When making their forecasts, investors use current and past realized volatility *as they perceive it*. If their perception of volatility is systematically biased by the after-effect, VIX will be distorted in specific ways, which we describe in Sections 2.1 and 2.2. Those predicted distortions are the focus of our empirical tests in Section 2.3. Our tests isolate the change in VIX that can be attributed to the after-effect specifically. We control for a number of potential confounds, including changes in fundamentals (proxied by market returns), changes in the variance risk premium (using lags of realized volatility, as per Baele et al. 2019), and microstructure effects.

We find a significant impact of the after-effect on the VIX. The effect presents itself consistently through an extensive set of robustness tests (which we report in Section 2.6). Strikingly, the magnitude of the distortions of VIX created by the after-effect is approximately the same as the impact of a 1% change in the S&P 500 index (the strongest predictor of changes in VIX).

To distinguish the after-effect from other explanations, we exploit two distinctive properties of the after-effect: its magnitude increases both with stimulus strength and with stimulus duration (e.g., Magnussen and Johnsen, 1986; Hershenson, 1989; Leopold et al., 2005). Therefore, if the VIX distortions found in our empirical tests are truly caused by the after-effect, the distortions should be maximal for regimes featuring extreme VH and VL levels, and nil for regimes in which the VH and VL levels do not depart markedly from the medium state. Moreover, the VIX distortions should also increase with regime duration (the exposure time to VH or VL levels). Our data support both predictions, while many alternative explanations are inconsistent with these features.

To address the most plausible remaining alternative explanations, in Section 2.5 we consider several kinds of expectation biases that have received solid empirical support in prior work, such as adaptive expectations,[4] anchoring, the gambler's fallacy,[5] and neglected risks.[6] We also account for investor learning about higher-order statistics, investor reactions to changes in jump risk, nonlinear changes in the variance risk premium, the possibility of short-term reversals in VIX, and Bayesian learning in a two-state regime switching model. We show that none of these explanations can predict the observed VIX distortions.

We further document that the VIX distortions attributable to the after-effect are asymmetric, mainly occurring following high volatility ("post-high"). One wonders whether this is chiefly due to the small number of very low "VL" volatility regimes in our data, which may have prevented us from identifying the after-effect following low volatility ("post-low"). Alternatively, the latter may be absent for equity volatility, due to VL levels around 7%, which is the mean VL value in our data, not being "very low" to the brain (it may fall in the medium category).

Therefore, in Section 3 of the paper, we return to the lab to answer this question of asymmetry. Our laboratory investigations allow us to study what the after-effect post-low may look like in markets with lower VL levels relative to those in the equity markets. In the bond market for example, it is typical to observe VL values around 4%, for treasury and investment-grade bonds in particular. To provide insights into these questions, we re-analyse the data from Payzan-LeNestour et al. (2016)'s study (henceforth, "the original condition") to separately measure the after-effect post-high and the after-effect post-low. We also re-run their laboratory experiment using parameter values for VL and VH that better reflect equity volatility (VL = 0.07; VH = 0.4), and exploit data available from a companion study (Payzan-LeNestour et al., 2021) in which the same experiment was run, this time with values for VL and VH that are more representative of the bond market (VL = 0.04; VH = 0.25–0.30). Pooling these data together gives us a stylized representation of the financial markets (henceforth, "the field condition").

The main findings are that the after-effect presents itself consistently across conditions, both post-high and post-low, with a stronger effect post-high. We also provide evidence that the after-effect post-low may be muted—possibly absent—for equity but significant for other kinds of assets.

The practical relevance of these findings is that they imply traders make systematic errors in assessing and responding to volatility following periods in which they adapt or become accustomed to a very low or very high level of volatility. We show that these perceptual errors result in inefficiencies in market prices of traded financial instruments such as equity options. A follow-up study shows that the inefficiencies are sufficiently large such that a trader that is aware of them can profitably exploit them by trading against the distortions in prices (Payzan-LeNestour et al., 2022).

1.1. Related literature

The VIX distortions caused by the after-effect add to the list of market "anomalies" documented in prior behavioural finance studies.[7] Also related to the current study, recent research shows how the stimulus to which a trader is exposed can systematically affect their decision making. For example, Borsboom et al. (2021) show

[4] For example, Malmendier and Nagel (2011), Greenwood and Shleifer (2014), Barberis et al. (2015).

[5] For example, Tversky and Kahneman (1971, 1974), Gilovich et al. (1985), and Chen et al. (2016).

[6] For example, Gennaioli et al. (2012, 2015), Goetzmann et al. (2016), and Gennaioli and Shleifer (2018).

[7] See, among others, De Long et al. (1990), Lee et al. (1991), Shleifer and Vishny (1997), Froot and Dabora (1999), Barberis and Shleifer (2002), Mitchel et al. (2002), and Lamont and Thaler (2003).

in controlled experiments that displaying financial market price charts at short horizons can lead to overtrading.

Closest to the present study, a growing body of work in psychology and experimental finance aims to measure risk perception separately from attitude towards perceived risk ("risk attitude"), following the lead of landmark psychology studies,[8] and in contrast to many risk-taking studies in which the two dimensions are confounded, resulting in lack of stability and poor predictive validity of behavioural experiments on risk taking (an issue well emphasized by Mata et al. 2018, amongst others). Decomposing risk taking into its two components has proved insightful in many respects. For example, evidence suggests that the documented domain specificity of risk taking arises from differences in perceived risk across domains, while risk attitude would be a stable strait (Weber and Milliman, 1997; Weber, 1998). Weber et al. (2013) provide evidence that risk taking changed after the Global Financial Crisis due to changes in risk perception, while risk attitude appeared to be stable around that period. Most recently, Huber et al. (2021) show that the recent COVID-19 shock decreased the level of perceived risk in investors, consistent with the idea of adaptive normalization proposed in this study. Weber et al. (2002) show that differences in perceived risk, rather than differences in risk attitude, explain gender differences in risk taking. Laudenbach et al. (2022) provide evidence that limited stock market participation arises from biased perceptions of downside risk. There is also ample evidence that the variance of a financial product is not a key determinant of investor risk perception relative to other aspects such as skewness and loss probability (e.g., Weber and Hsee, 1988; Klos et al., 2005, Unser, 2000; Huber et al., 2019; Holzmeister et al., 2020; Zeisberger, 2022). Contrary to these studies, the current study does not ask how the average investor operationalizes risk. Rather, we ask how those who need to properly perceive volatility for their decision-making (e.g., professional traders) assess it. In light of the foregoing studies, we caution against extrapolating the current insights into volatility perception to the topic of risk perception in a broad sense, since "risk" is a multidimensional variable with different meanings for different actors.

2. Field study

2.1. Identifying volatility regimes that may induce after-effects

To test for biased volatility perception in the field, we use data on the S&P 500 index and VIX index values for the period January 2, 1996 to December 31, 2020 (the VIX measures implied volatility in the S&P 500 index).[9] Denote the log S&P 500 index value by p. A daily interval $[t-1,\ t]$ consists of N tick-by-tick observations $\{t_0, t_1, \ldots t_N\}$. We first compute the realized variance at the optimal sampling frequency K[10] as $RV^2_{t-1,t} = \frac{1}{K}\sum_{i=0}^{N-K}[p(t_{i+K}) - p(t_i)]^2$. Realized volatility for the daily interval $[t-1,\ t]$ is computed by taking the square root of $RV^2_{t-1,t}$ and annualizing using a year of 252 business days: $RV_{t-1,t} = \sqrt{RV^2_{t-1,t} \times 252}$. To simplify notation, we refer to realized variance and realized volatility for the daily interval $[t-1,\ t]$ with a single time subscript corresponding to the end of the daily interval, RV^2_t and RV_t (t indexes business days).

To identify volatility regimes in our data, we must first identify very high, very low, and neutral volatility states. We use the approximate normality of the log realized volatility to create volatility bins.[11] We compute the mean and standard deviation of the distribution of the daily log realized volatility, during a rolling three-month (63 business days) window. We then define a very high (VH) volatility level as one that is more than x standard deviations above the mean, and a very low (VL) volatility level as one that is more than x standard deviations below the mean. A medium or neutral volatility level (M) is one that falls within y standard deviations of the mean.[12]

We then create volatility regime indicator variables defined over a four-day period. The choice of four days is motivated by empirical tradeoffs between sufficient statistical power and the need to minimize effects from market microstructure.[13] In Section 2.3 we show that the after-effect is stronger for a longer volatility regime. $VolReg^+_t$ takes the value of 1 if we observe very high volatility levels in the three preceding days ("high volatility state") and a neutral volatility level on day t (zero otherwise). $VolReg^-_t$ takes the value -1 if we observe very low volatility levels in the three preceding days ("low volatility state") and a neutral volatility level on day t (zero otherwise). We also define a combined measure $VolReg_t = VolReg^+_t + VolReg^-_t$ that is $+1$ following a transition from very high to neural volatility and -1 following a transition from very low to neural volatility:

$$VolReg^+_t = \begin{cases} +1 \ if\ \{LnRV_{t-3},\ LnRV_{t-2},\ LnRV_{t-1}, LnRV_t\} = \{VH,\ VH, VH, M\} \\ 0 \ \ otherwise \end{cases}$$

$$VolReg^-_t = \begin{cases} -1 \ if\ \{LnRV_{t-3},\ LnRV_{t-2},\ LnRV_{t-1}, LnRV_t\} = \{VL,\ VL, VL, M\} \\ 0 \ \ otherwise \end{cases}.$$

$$VolReg_t = \begin{cases} +1 \ if\ \{LnRV_{t-3},\ LnRV_{t-2},\ LnRV_{t-1}, LnRV_t\} = \{VH,\ VH, VH, M\} \\ -1 \ if\ \{LnRV_{t-3},\ LnRV_{t-2},\ LnRV_{t-1}, LnRV_t\} = \{VL,\ VL, VL, M\} \\ 0 \ \ otherwise \end{cases} \tag{1}$$

where $LnRV_t$ is log realized volatility on day t.[14] The identification of volatility states (very high, very low, and neutral) and regimes (transitions from prolonged very high or very low volatility to neutral volatility) is illustrated in Fig. 1.

This figure illustrates how volatility regimes are defined. A very-high-to-neutral transition ($VolReg^+_t = 1$) occurs when realized volatility is very high (greater than x standard deviations above the mean) for at least three consecutive days and then neutral (within y standard deviations from the mean) the next day. Similarly, a very-low-to-neutral transition ($VolReg^-_t = -1$) occurs when

[8] For example, Sitkin and Weingart (1995), Weber and Milliman (1997), and Slovic (1997).

[9] The sample starts in 1996 as that is the first year available in the Refinitiv (formerly Thomson Reuters) dataset that we use for market data and the sample end reflects the latest available data at the time of analysis.

[10] The optimal sampling frequency depends on the degree of trading activity, among other factors, which changes substantially through time from the start to the end of our 18-year sample. We therefore use three different sampling frequencies (ten, five and three minutes) in three different time periods, increasing the frequency in line with trading activity.

[11] An additional reason for working with log realized volatility is that its volatility shows little persistence (Corsi et al., 2008).

[12] We choose to use symmetrical intervals because the distribution is approximately symmetrical (see the Internet Appendix). We use five buckets to avoid threshold effects happening when volatility has been in the highest or lowest bucket and a small change brings it into the adjacent middle bucket. Setting $x = y$ collapses the five buckets into three adjacent ones. Our results are robust to different choices of the look-back window (see Internet Appendix).

[13] Specifically, if we extended the length of volatility regimes to weeks or months to capture longer-horizon after-effects, we would have too few observations to have powerful tests. Conversely, if we go into intra-day horizons to capture higher-frequency after-effects, we run into a range of potentially confounding effects from market microstructure (bid-ask bounce and so forth).

[14] Note that volatility regimes ($VolReg_t = \pm 1$) can involve very high or very low realized volatility that persists for more than three days before transitioning to the neutral level. Below we investigate the impact of the length of the stimuli on effect size.

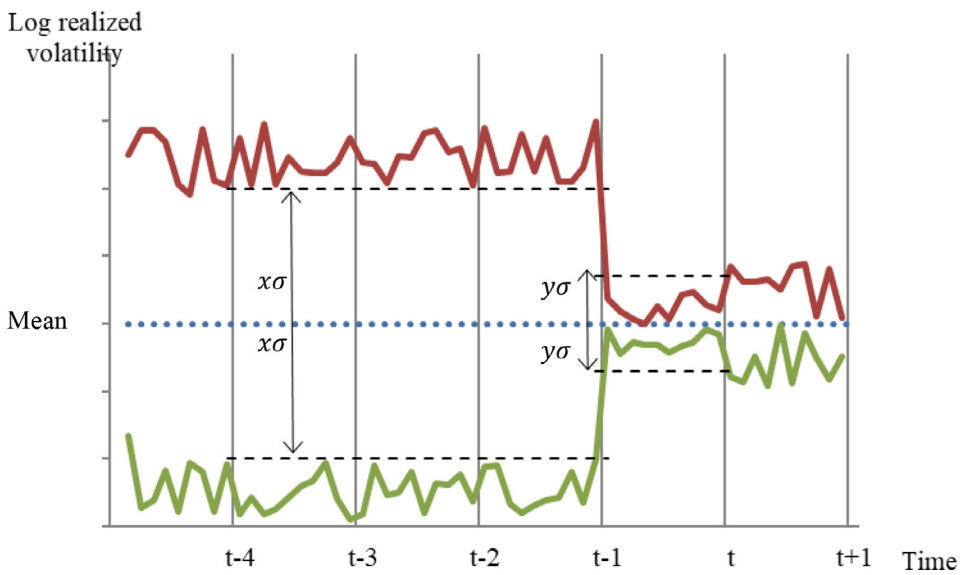

Fig. 1. Method used to identify the volatility regimes.

realized volatility is very low (more than x standard deviations below the mean) for at least three consecutive days and then neutral (within y standard deviations from the mean) the next day.

The higher we set x and the lower we set y the larger the differences between the stimulus in the adaptation phase (the very high or very low volatility) and the neutral volatility in the transition phase and therefore the stronger the expected after-effect. The downside of setting a very high x and low y is that we have fewer transitions in the data—fewer opportunities to test for an after-effect.[15]

After removing the first three months of the sample used in the rolling window that determines high/low/neutral levels, we are left with 6195 daily observations. When $x = y = 1$, there are about 200 volatility regimes (transitions from very high or very low volatility states to the neutral state) over the whole sample 1996–2020.[16] The fairly small number of observations works against us finding a significant result and is therefore a conservative feature of our empirical tests. Fig. 2 illustrates the temporal distribution of volatility regimes for $x = y = 1$. Regimes occur regularly throughout the sample with some evidence of clustering, for example, during the second semester of 2009.

The horizontal axis measures time from the start of our sample (April 3, 1996) until the end (December 31, 2020). Vertical lines to $+1$ indicate very-high-to-neutral transitions in realized volatility ($VolReg_t^+ = +1$) and vertical lines to -1 indicate very-low-to-neutral transitions ($VolReg_t^- = -1$).

Table 1 Panel A reports the number of regimes for a range of x and y between 1.00 and 1.75 standard deviations. Table 1 Panel B reports the average absolute difference between the log realized volatility in neutral states and the log realized volatility in very-high and very-low states for volatility regimes. The absolute difference in volatility levels increases with the threshold that defines an extreme volatility level (x) and it decreases with the threshold that defines a neutral volatility level (y). The average jump from either a very high or very low volatility state to a neutral volatility state is 0.42 in log terms (about 40% in realized volatility terms).

2.2. Structural model for empirical tests

According to the theory, an after-effect occurs when transitioning from a prolonged very low or very high volatility level to a neutral volatility level (neither high nor low). To test the theory, we investigate the change in perceived volatility during these transitions. We extract the changes in perceived volatility from the VIX, a model-free measure of volatility expectations, using a structural model set out below. In the baseline version of this model, without loss of generality, we assume rational expectations (assuming adaptive expectations instead strengthens our main conclusions, more on this below).

The structural model begins by expressing VIX squared (which is the price of a synthetic variance swap)[17] as the sum of expected realized variance and a variance risk premium (like in Carr and Wu 2006 and Bollerslev et al. 2009):

$$VIX_t^2 = \mathbb{E}_t\left[RV_{t,t+22}^2\right] + VRP_t, \tag{2}$$

where $\mathbb{E}_t$ is the expectation under the statistical probability measure and VRP_t is the variance risk premium.[18]

[15] In our empirical tests we balance these competing considerations. In doing so, we discard choices of x and y that do not yield a sufficiently large number of regimes (from a statistical viewpoint). We show that our results are robust to different choices of x and y and in fact we exploit different combinations of x and y to test how the strength of the after-effect varies with the strength of the stimulus.

[16] The relatively low number of regime change is not surprising: our regime indicator is defined over four days and we have 6,195 days in the sample. So we have a maximum of 1,549 non-zero values. To get a non-zero value, realized volatility has to stay in the tail of the distribution for three days in a row before jumping. This is an unlikely path (albeit it is possible given the persistent nature of realized volatility and the presence of jumps).

[17] VIX is quoted as an annualized standard deviation (volatility), so VIX squared is annualized variance. VIX squared corresponds to variance over the following 30 calendar days, which for expositional ease, we denote as 22 business days to avoid introducing a second measure of time.

[18] Carr and Wu (2006) find that VIX is on average around five percentage points higher than realized volatility as a result of the variance risk premium. The variance risk premium implies that investors require compensation for bearing variance risk (a position that incurs losses when variance is unexpectedly high, e.g., the short side of a variance swap). Equivalently, investors are willing to pay a positive premium (accept a negative expected return) to hedge variance risk with a contract that has a positive (negative) payoff when volatility is unexpectedly high (low). See the Internet Appendix for more details. As noted by Carr and Wu (2006), VIX is a discretized approximation of the variance. The approximation is exact in the absence of jumps in the underlying. When the underlying index exhibits jumps, the approximation error is negligible (Carr and Wu, 2009; Jiang and Tian, 2007).

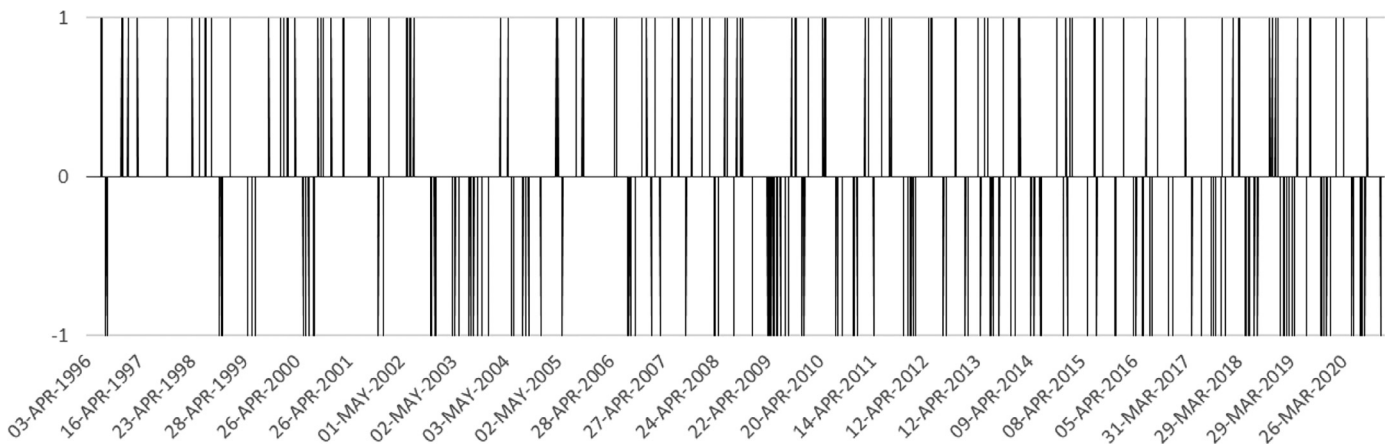

Fig. 2. Distribution of the regime changes in realized volatility through time.

Table 1
Number and strength of volatility regimes for different threshold values
Panel A reports the number of realized volatility "regimes" (very-high-to-neutral ($VolReg_t^+ = 1$) in the columns to the left and very-low-to-neutral ($VolReg_t^- = -1$) in the columns to the right) for different threshold values. Panel B reports the average absolute difference between the log realized volatility in neutral states and the log realized volatility in very-high and very-low states for volatility regimes. Columns report different values of the threshold that defines very high and very low volatility states (volatility that is greater than x standard deviation from the mean). Rows report different values of the threshold that defines the neutral volatility state (volatility that is within y standard deviations from the mean).

	Very-high-to-neutral regimes ($VolReg_t^+ = 1$)				Very-low-to-neutral regimes ($VolReg_t^- = -1$)			
Panel A: Number of regimes								
	x				x			
y	1.00	1.25	1.50	1.75	1.00	1.25	1.50	1.75
1.00	105	52	25	10	135	60	23	5
1.25	.	78	44	24	.	84	30	7
1.50	.	.	59	35	.	.	42	8
1.75	.	.	.	40	.	.	.	15
Panel B: Volatility differences between extreme and neutral states								
	x				x			
y	1.00	1.25	1.50	1.75	1.00	1.25	1.50	1.75
1.00	0.37	0.40	0.44	0.50	0.29	0.32	0.36	0.45
1.25	.	0.36	0.40	0.42	.	0.28	0.31	0.39
1.50	.	.	0.37	0.38	.	.	0.25	0.37
1.75	.	.	.	0.36	.	.	.	0.30

By differencing (2), we obtain:

$$\Delta VIX_t^2 = VIX_t^2 - VIX_{t-1}^2 = \mathbb{E}_t\left[RV_{t,t+22}^2\right] - \mathbb{E}_{t-1}\left[RV_{t-1,t+21}^2\right] + \Delta VRP_t. \quad (3)$$

Daily changes in expected future volatility are driven by changes in daily realized volatility. We demonstrate this in the Internet Appendix, but the intuition is simply as follows. Forecasts of future volatility are based on current and past realized volatility. So, as we move from one day to the next, agents adjust their forecasts of future volatility using the new information that they have, which is that day's realized volatility. Therefore, (3) becomes:

$$\Delta VIX_t^2 \approx \beta_D\ \Delta RV_t^2 + \Delta VRP_t. \quad (4)$$

If agents' perceptions of volatility are distorted due to after-effects by a perception error, PE, then we have the following (using subscript π for perceived volatility):

$$RV_{\pi,t}^2 = \begin{cases} RV_t^2 - PE,\ after\ a\ prolonged\ period\ of\ high\ volatility \\ RV_t^2 + PE,\ after\ a\ prolonged\ period\ of\ low\ volatility \\ RV_t^2,\ otherwise. \end{cases}$$
$$= RV_t^2 - PE.VolReg_t^+ - PE.VolReg_t^-$$
$$= RV_t^2 - PE.VolReg_t \quad (5)$$

Taking first differences gives:

$$\Delta RV_{\pi,t}^2 = \Delta RV_t^2 - PE.VolReg_t^+ - PE.VolReg_t^-$$

$$\Delta RV_{\pi,t}^2 = \Delta RV_t^2 - PE.VolReg_t \quad (6)$$

Allowing for perception error in (4) by replacing the change in realized volatility with the perceived change in volatility (6), we get:

$$\Delta VIX_t^2 \approx \beta_D\ \Delta RV_t^2 - \beta_D PE.VolReg_t^+ - \beta_D PE.VolReg_t^- + \Delta VRP_t.$$

$$\Delta VIX_t^2 \approx \beta_D\ \Delta RV_t^2 - \beta_D PE.VolReg_t + \Delta VRP_t. \quad (7)$$

At daily frequencies, changes in the variance risk premium are negligible. As Merton (1980) and Bollerslev et al. (2011) highlight, the variance risk premium is slow-moving, highly persistent and varies with the business cycle, rather than day to day. Therefore, the equation above shows that daily changes in VIX are primarily driven by daily changes in realized volatility, as well as a distortion due to errors in volatility perception. After-effects theory predicts that those distortions occur at specific times following the volatility regimes captured by the $VolReg_t$ variable or its components, $VolReg_t^+$ and $VolReg_t^-$. Therefore, if $VolReg_t$ explains changes in VIX after controlling for other factors such as changes in realized

Table 2
VIX distortions induced by the volatility after-effect
This table reports coefficient estimates from Regressions (8–11) where the dependant variable is the daily change in log VIX. Negative values of the coefficient of $VolReg_t^+$ or $VolReg_t^-$ (the indicator variables for high-to-neural and low-to-neutral volatility regimes) in Columns (1–4 and 7) are consistent with VIX distortions caused by the volatility after-effect phenomenon. Negative values of the coefficients of $VolReg_t^+.x_t$ (Column (5)), $VolReg_t^-.x_t$ (Column (8)), $VolReg_t^+.z_t$ (Column (6)), and $VolReg_t^-.z_t$ (Column (9)) are consistent with a positive relationship between respectively stimulus strength and after-effect magnitude, and stimulus duration and after-effect magnitude. x_t is a measure of how extreme (in terms of number of standard deviations from the mean) realized volatility has been on average during the adaptation phase. z_t is the duration of the adaptation phase, measured by the number of days the phase lasts, minus three. The other coefficients are the coefficients of the control variables: $\Delta LnRV_t$ and $\Delta LnRV_{t-1}$ are daily log realized volatility and its lag. r_t is the market return during day t and $r_t^- = \min(r_t, 0)$. $\Delta LnVIX_{t-1}$ is the lagged daily change in log VIX. All the regressions use the threshold parameters $x = 1.75$ and $y = 1.50$ except for Columns (5–6 and 8–9) which uses threshold parameters $x = 1.00$ and $y = 1.00$ (because they test the effects of stimulus strength and duration from a weak to strong level). t-statistics (in parenthesis) are calculated with heteroskedasticity robust standard deviations. ***, **, and * denote coefficients significant at the 1%, 5%, and 10% level respectively.

	(1)$\Delta LnVIX_t$	(2)$\Delta LnVIX_t$	(3)$\Delta LnVIX_t$	(4)$\Delta LnVIX_t$	(5)$\Delta LnVIX_t$	(6)$\Delta LnVIX_t$	(7)$\Delta LnVIX_t$	(8)$\Delta LnVIX_t$	(9)$\Delta LnVIX_t$
$VolReg_t^+$	−3.844***	−2.806***	−2.765***	−2.628***	−0.743	−1.568***			
	(−3.10)	(−2.82)	(−2.85)	(−2.72)	(−0.82)	(−2.64)			
$VolReg_t^+.x_t$					−2.352**				
					(−2.00)				
x_t					−0.123				
					(−0.90)				
$VolReg_t^+.z_t$						−0.658**			
						(−2.02)			
$VolReg_t^-$							−0.645	−0.253	−0.298
							(−0.97)	(−0.27)	(−0.45)
$VolReg_t^-.x_t$								−2.718	
								(−1.34)	
x_t								−0.307**	
								(−2.33)	
$VolReg_t^-.z_t$									−0.436
									(−0.79)
$\Delta LnRV_t$	0.075***	0.030***	0.040***	0.043***	0.042***	0.043***	0.044***	0.044***	0.044***
	(19.38)	(11.42)	(11.69)	(12.01)	(11.61)	(11.89)	(12.15)	(12.14)	(12.36)
$\Delta LnRV_{t-1}$			0.017***	0.020***	0.020***	0.020***	0.021***	0.021***	0.021***
			(5.42)	(6.17)	(5.96)	(6.03)	(6.27)	(6.20)	(6.33)
r_t		−3.514***	−3.420***	−3.429***	−3.404***	−3.407***	−3.449***	−3.426***	−3.446***
		(−18.76)	(−18.38)	(−18.62)	(−18.71)	(−18.73)	(−18.85)	(−18.69)	(−18.82)
r_t^-		1.016***	1.024***	0.965***	1.001***	0.994***	0.934***	0.980***	0.941***
		(3.21)	(3.27)	(3.10)	(3.25)	(3.21)	(3.02)	(3.18)	(3.04)
$\Delta LnVIX_{t-1}$			−0.116***	−0.115***	−0.115***	−0.116***	−0.115***	−0.114***	−0.115***
			(−4.72)	(−4.66)	(−4.68)	(−4.71)	(−4.65)	(−4.62)	(−4.67)
Day of The Week FE				YES	YES	YES	YES	YES	YES
R^2	9.09%	55.41%	56.61%	57.71%	57.86%	57.85%	57.64%	57.69%	57.64%
Observations	6195	6195	6195	6195	6195	6192	6195	6195	6192

volatility, there is evidence of perceptual errors that are consistent with after-effects.

The structural model above implies that the presence of after-effects in volatility perceptions can be identified by testing whether $VolReg_t^+$ or $VolReg_t^-$ explain changes in VIX. In our empirical tests, we use a log specification of (7) for consistency with the definition of the variables $VolReg_t^+$ and $VolReg_t^-$.[19] The log specification is not pivotal for our results and in robustness tests we find similar results, in some cases even stronger, under alternative specifications including specifications that are not in logs. The benchmark form of our regression is thus :[20]

$$\Delta LnVIX_t = \alpha + \beta_1\ VolReg_t^+ + \beta_2\ VolReg_t^- + \gamma\ \Delta LnRV_t + \varepsilon_t. \quad (8)$$

The main coefficients of interest are β_1 and β_2: if volatility after-effects distort VIX as described by our structural model, β_1 and β_2 should be significantly negative.

2.3. Results measuring the impact of a volatility regime on VIX

Table 2 Column (1) reports estimates from the baseline Regression (8) starting with the tests of very-high-to-neutral volatility transitions. The impact of $VolReg_t^+$ on changes in VIX has a negative sign, which is consistent with the presence of volatility after-effects following very high volatility, and it is statistically significant. The economic impact of $VolReg_t^+$ is large: a transition from a very high volatility state to neutral volatility changes VIX by about 3.8%.

Note that our benchmark regression assumes the ΔVRP_t term in (7) is negligible following Merton (1980) and Bollerslev et al. (2011). To ensure that such neglect of ΔVRP_t is not pivotal for our findings, we augment the baseline Regression (8) with variables that control for ΔVRP_t. We control for the first lag of changes in realized volatility (the results are robust to adding more lags), in line with Baele et al. (2019), whose findings suggest that volatility is the main predictor of variations in the variance risk premium. We also include the market return during the transition period (r_t) as a further proxy for changes in the variance risk premium, as well as negative market returns ($r_t^- = \min(r_t, 0)$) to account for possible leverage effects. To account for any possible auto-correlation of changes in VIX, we include the first lag of changes in VIX. This first lag accounts for any short-term reversals in VIX, like the short-term reversals often seen in equity returns due to illiquidity. Finally, we include dummy variables for the "day-of-the-week effect" in VIX.[21] The

[19] We multiply the log difference by 100 to make it consistent with the definition of S&P 500 returns.

[20] In logging the series, the squares become linear terms, with the factor of two being absorbed into the corresponding coefficients and regression intercept.

[21] Fleming et al. (1995) provide evidence of intra-week seasonality in VIX, with one explanation being that the weekend creates a different level of implied volatil-

regression with the complete set of control variables is thus:

$$\Delta LnVIX_t = \alpha + \beta_1\ VolReg_t^+ + \beta_2\ VolReg_t^- + \gamma_0\ \Delta LnRV_t + \gamma_1\ \Delta LnRV_{t-1}$$

$$+\delta\ r_t + \delta^-\ r_t^- + \rho_1\ \Delta LnVIX_{t-1} + \sum_{i=2}^{5}\theta_i D_{it} + \varepsilon_t, \qquad (9)$$

where $\{D_{it}\}_{i=2,3,4,5}$ are dummy variables for Tuesday to Friday (Monday is base case).

The regressions with the full set of control variables are also reported in Table 2, as Columns (2–4). The coefficients of the key variable, $VolReg_t^+$, is not overly affected by the additional control variables and remains remarkably stable across all regressions.

Turning to very-low-to-neutral volatility transitions and going straight to models with the control variables, Table 2 Column (7) shows that the impact of $VolReg_t^-$ on changes in VIX also has a negative sign, which is consistent with the presence of volatility after-effects following very low volatility, but the magnitude of this effect is smaller and is not statistically significant. We explore this asymmetry in the after-effect in the laboratory tests later in the paper to better understand its cause.

In the Internet Appendix (section A.7), we report results that show strong evidence of after-effects (across all our key empirical tests) when using the combined measure, $VolReg_t$ instead of $VolReg_t^+$ and $VolReg_t^-$.

In the foregoing regressions we use threshold parameter $x = 1.75$ and $y = 1.50$ (recall x determines the volatility level during the adaptation phase and y determines the volatility level in the neutral state). This parameter choice for x and y is to ensure both a sufficiently large difference between the volatility states and a sufficiently large number of regimes. Importantly, our results are robust to using different parameter choices for x and y. Table 3 reports the estimated coefficients and significance of $VolReg_t^+$ and $VolReg_t^-$ in Regression (9) for different values of both parameters. In all cases the estimated coefficient of $VolReg_t^+$ is negative and statistically significant as per the previous regression results, whereas $VolReg_t^-$ is negative throughout but only significant in one case.[22] The largest coefficient of $VolReg_t^+$ is 4.315 and the smallest is 2.483, which implies that the impact of the after-effect on VIX following high volatility is similar in magnitude to the impact of a 1% change in S&P 500.

Our baseline results so far measure the VIX distortion in a one-day period following the volatility transition from very high or very low. We also test whether the VIX distortions persist beyond the day that the after-effect is triggered. To do this, we time shift the left-hand side of Regression (9) so that the impact on VIX is measured at time $t+1$ rather than time t. We find that about one-third (approximately 33%) of the distortion attributable to after-effects following high volatility is reversed the following day.[23] Thus, a substantial fraction of the effect wears off within day, but some of the distortion persists beyond one day. The tendency for the after-effect to wear off is consistent with prior studies and the fact that the after-effect is a temporary perceptual distortion while the brain adapts to the new level of the stimulus. In a practical sense, it implies the asset price distortions are short-lived. Despite being temporary, a follow-up paper shows that they are sufficiently long-lived that they present profitable arbitrage opportunities (Payzan-LeNestour et al., 2022).

Table 3
Strength of the volatility after-effect for different volatility thresholds
This table reports coefficients for the $VolReg_t^+$ (Panel A) and $VolReg_t^-$ (Panel B) variables, which measure very-high-to-neutral and very-low-to-neutral volatility transitions, respectively. The coefficient estimates are obtained from Regression (9). Columns report different values of the threshold that defines very high and very low volatility states (volatility that is greater than x standard deviation from the mean). Rows report different values of the threshold that defines the neutral volatility state (volatility that is within y standard deviations from the mean). t-statistics (in parenthesis) are calculated with heteroskedasticity-robust standard errors. ***, **, and * denote coefficients significant at the 1%, 5%, and 10% level respectively.

	x			
y	1.00	1.25	1.50	1.75
Panel A: Transitions from VH to M				
1.00	−2.493***	−3.203***	−4.315***	−2.953**.
	(−5.34)	(−4.88)	(−3.90)	(−2.23)
1.25	.	−2.920***	−4.122***	−3.413***
		(−5.03)	(−5.51)	(−4.04)
1.50	.	.	−2.483***	−2.628***
			(−2.68)	(−2.72)
1.75	.	.	.	−2.990***
				(−3.40)
Panel B: Transitions from VL to M				
1.00	−0.268	−0.536	−1.615*	.
	(−0.82)	(−1.03)	(−1.67)	
1.25	.	−0.419	−1.173	.
		(−1.00)	(−1.49)	
1.50	.	.	−0.645	.
			(−0.97)	
1.75	.	.	.	−0.699
				(−0.91)

We also examine whether the strength of the after-effect following high volatility is diminished when the adaptation phase of the volatility regime spans a weekend. We do this by adding an interaction of the $VolReg_t^+$ variable with a dummy variable for whether the adaptation phase spans a weekend. We find that after-effects are about 18% stronger if the adaptation phase does not include a weekend.[24] Therefore, our main results that include volatility regimes that have weekends in the adaptation phase, tend to understate the magnitudes of after-effects that arise following three continuous days of high or low volatility exposure.

2.4. After-effect magnitude as a function of regime strength and regime duration

To further put our theory to the test, we exploit the fact that it makes two very specific predictions: (i) the more extreme the volatility during the adaptation period (the time spent in a very high or very low volatility state), the stronger the neuronal adaptation and hence the larger the after-effect in the neutral state; and (ii) the longer the adaptation period, the stronger the neuronal adaptation and hence the larger the after-effect. Our data support both predictions.

We find that the more extreme the volatility levels during the three days preceding a transition to neutral volatility, the larger the coefficient on the $VolReg_t^+$ variable and the larger the distortion in VIX (we focus on the transitions from high volatility because our baseline results show that is where a significant after-effect is present). This result is apparent in Fig. 3, which displays the coefficient of $VolReg_t^+$ as a function of the level of (log) real-

ity on Friday compared to other weekdays. We account for his effect by controlling for the day of the week.

[22] The coefficients of the control variables are virtually unchanged for the different values of x and y.

[23] These additional results are reported in the Internet Appendix Table A.20. The coefficient of $VolReg_t^+$ decreases in magnitude from −2.628 in Table 2 to −1.770 when the left-hand side variable is time shifted, implying a one-third reversal of the effect within one day.

[24] These additional results are reported in the Internet Appendix Table A.21. The coefficient of $VolReg_t^+$ increases in magnitude from −2.628 in Table 2 to −3.102 when the regime does not include a weekend, implying a 18% increase in the after-effect strength.

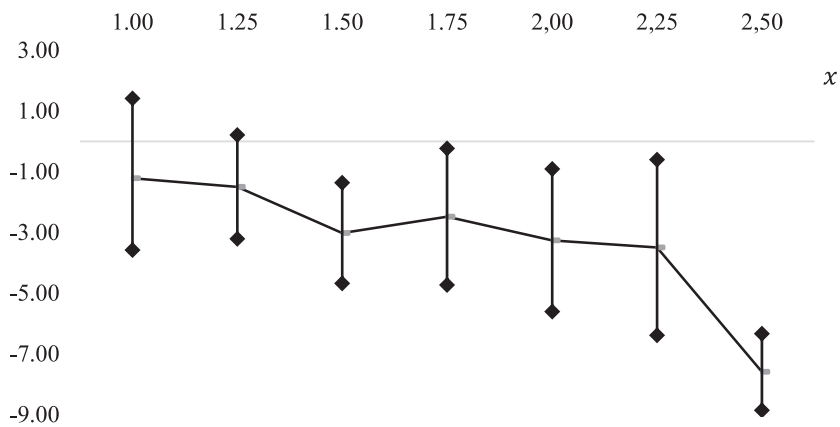

Fig. 3. Stimulus strength and magnitude of the volatility after-effect.

ized volatility in the adaptation phase. The size of the after-effect steadily increases with the average level of realized volatility in the adaptation phase, as predicted by the theory.

This figure plots the coefficient of $VolReg_t^+$ (vertical axis) in Regression (9) for different values of x. Negative values of the coefficient of $VolReg_t^+$ are consistent with a perceptual bias due to after-effects. Panel A is a plot of the coefficient estimates (strength of the after-effect) and 95% confidence intervals, for the different values of x, the threshold that defines very high and very low volatility states (horizontal axis).

The estimates in Fig. 3 allow for non-linearity in the relation between the strength of the volatility in the adaptation phase and the strength of the resulting after-effect. As an alternative way to assess the relationship between stimulus strength and after-effect magnitude, we augment our baseline Regression (9) with an interaction $VolReg_t^+.x_t$, where x_t measures how extreme (in terms of number of standard deviations from the mean) realized volatility has been on average during the past three-day adaptation phase. We also control for the level of x_t as a standalone term:

$$\Delta LnVIX_t = \alpha + \beta_1 VolReg_t^+ + \beta_2\ VolReg_t^+.x_t + \beta_3 x_t + \gamma_0\ \Delta LnRV_t$$

$$+\gamma_1\ \Delta LnRV_{t-1} + \delta\ r_t + \delta^-\ r_t^- + \rho_1\ \Delta LnVIX_{t-1} + \sum_{i=2}^{5}\theta_i D_{it} + \varepsilon_t. \tag{10}$$

Table 2 Column (5) reports the results from Regression (10). The main coefficient of interest, β_2, is significantly negative, indicating a statistically significant positive relation between stimulus strength and after-effect magnitude as predicted by theory. The implied marginal effects from this linear specification are similar to those of the non-linear tests reported in Fig. 3. For example, at a value of $x_t = 2$ the estimated after-effect following high volatility according to Regression (10) is -5.447, while at $x_t = 3$ the estimated after-effect is -7.799 and these values are similar to those in Fig. 3.

To test prediction (ii), we modify $VolReg_t^+$ so that the adaptation period spans three, four, or five days.[25] The after-effect increases with the number of days in the adaptation window, as the theory predicts. This finding is illustrated in Fig. 4, which plots the coefficient on $VolReg_t^+$ for different length adaptation windows. The absolute value of the estimated coefficient of the $VolReg_t^+$ variable increases for longer adaptation phases, suggesting a stronger after-effect for longer exposure to the volatility stimulus.

This figure plots the coefficient of $VolReg_t^+$ (vertical axis) in Regression (9) for different values of the stimulus duration. Negative values of the coefficient of $VolReg_t^+$ are consistent with a perceptual bias due to after-effects. The figure plots the coefficient estimates (strength of the after-effect) and 95% confidence intervals, for three different values of the stimulus duration (the period of very high volatility, horizontal axis).

Again, as an alternative way to test the relation between stimulus duration and after-effect magnitude, we augment our baseline Regression (9) with the interaction term $VolReg_t.z_t$, where z_t is the duration of the adaptation phase, measured by the number of days the phase lasts, minus three. Recall that three days is the minimum length for the adaptation phases in our tests and thus z_t is zero for regimes with three-day adaptation phases and increases by one for every additional day of extreme volatility in the adaptation phase.

$$\Delta LnVIX_t = \alpha + \beta_1 VolReg_t^+ + \beta_2\ VolReg_t^+.z_t + \gamma_0\ \Delta LnRV_t + \gamma_1\ \Delta LnRV_{t-1}$$

$$+\delta\ r_t + \delta^-\ r_t^- + \rho_1\ \Delta LnVIX_{t-1} + \sum_{i=2}^{5}\theta_i D_{it} + \varepsilon_t. \tag{11}$$

Table 2 Column (6) reports the results from Regression (11).[26] The main coefficient of interest, β_2, is significantly negative, indicating a statistically significant positive relation between stimulus duration and after-effect magnitude as predicted by theory.

The results in this subsection strengthen the evidence for volatility after-effect theory as the mechanism responsible for the VIX distortions by showing that the data are consistent with two very specific predictions made by the theory. Moreover, our finding that the VIX distortions vanish as the length of the adaptation window shrinks shows that those distortions are not caused by jumps in volatility per se. Only jumps that are preceded by prolonged exposure to very high or very low volatility cause the effect, as predicted by volatility after-effect theory.

2.5. Testing competing explanations

In this subsection, we report further tests of competing explanations for the results. These include adaptive expectations about

[25] In the Internet Appendix, we document that there are fewer "transitions" using a three-day adaptation window than a two-day window and even fewer when using longer windows (as one would expect). The absolute differences between the volatility level in the very high / very low state compared to the neutral state are almost identical when using the two-day and three-day adaptation windows.

[26] In Regression (11) we do not include z (the number of days in the adaptation phase) separately because it can only be measured when the $VolReg_t$ variables not equal to zero. Thus, if it was included as a stand-alone term, it would simply be colinear with the interaction term.

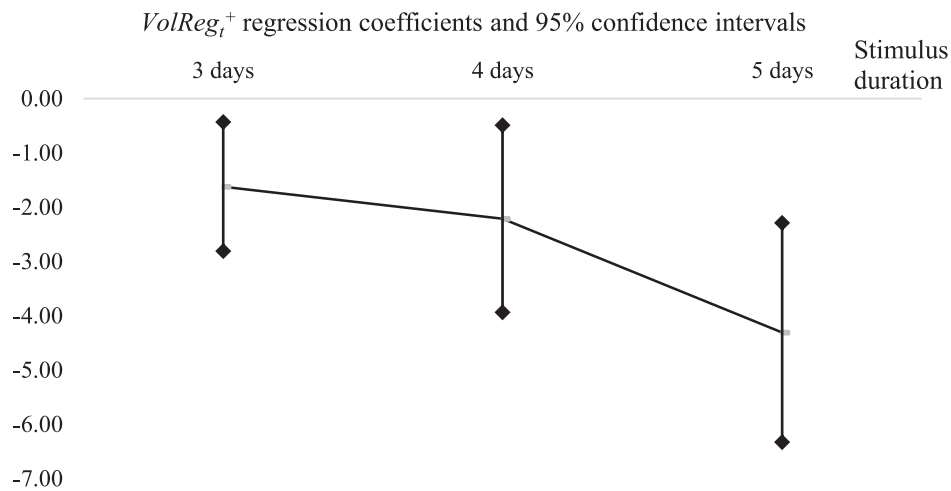

Fig. 4. Stimulus duration and magnitude of the volatility after-effect.

volatility changes, jump risk, non-linearity in the variance risk premium, and learning about kurtosis.

First, we consider the possibility that agents hold adaptive/extrapolative expectations about volatility *changes*, whereby after seeing an increase (resp. decrease) in realized volatility, they expect a further increase (resp. decrease). Adaptive expectations prevail in many domains.[27] According to the adaptive expectations hypothesis, immediately after transitioning from a high (resp. low) volatility state to a neutral state, the agent expects volatility to further decrease (resp. increase). Consequently, the agent revises his expectation of 30-day future volatility downward (resp. upward), causing a negative (resp. positive) change in VIX. Our finding that $\beta < 0$ could reflect this, rather than the volatility after-effect phenomenon predicted by volatility after-effect theory.

To tease apart the after-effect and adaptive-expectations-about-volatility-changes theories, we construct a "placebo" test in which we slightly modify our volatility regime indicator variable. The modified indicator, which we denote as $ModVolReg_t$, measures jumps between adjacent volatility states after a period of stability in volatility levels. That is, under the $ModVolReg_t$ measure, realized volatility jumps from a prolonged state of neutral volatility to a very high or very low volatility state:

$$ModVolReg_t = \begin{cases} +1 \ if \ \{LnRV_{t-3}, \ LnRV_{t-2}, \ LnRV_{t-1}, LnRV_t\} = \{M, \ M, M, VH\} \\ -1 \ if \ \{LnRV_{t-3}, \ LnRV_{t-2}, \ LnRV_{t-1}, LnRV_t\} = \{M, \ M, M, VL\} \\ 0 \ otherwise \end{cases} . \tag{12}$$

The jumps captured by the modified volatility regime variable $ModVolReg_t$ do *not* cause an after-effect, however they do cause the foregoing expectation bias, so the value of the estimated coefficient of $ModVolReg_t$ in Regression (9) (replacing the $VolReg_t$ variables with $ModVolReg_t$ in the regression) is a test of the two theories against one another: if volatility after-effect theory is correct, the coefficient should not be significantly different from zero (to reflect the absence of after-effect induced by the jumps captured by $ModVolReg_t$), whereas if our main findings are chiefly driven by adaptive expectations about volatility changes, the coefficient should be significantly different from zero.

Table 4 shows the data favour volatility after-effect theory against the expectation bias hypothesis. There is no effect for transitions from a neutral state to a very high or very low volatility state: the coefficient of the $ModVolReg_t$ variable is small and not statistically significant.[28]

By showing that it is not the occurrence of jumps in volatility per se that drives the bias in VIX that we document, but only those jumps that induce an after-effect (namely, transitions from prolonged extreme volatility to neutral volatility), the evidence in Table 4 also rules out other plausible explanations that are based on large jumps in volatility. For example, changes in the variance risk premium, which could be non-linear with respect to the change in variance, or the possibility that large volatility jumps lead agents to revise their beliefs about kurtosis, thereby impacting options prices.

Table 4
Placebo test
This table reports coefficient estimates for the modified measure of volatility regimes, $ModVolReg_t$ (which measures transitions from the neutral volatility state to very high and very low volatility states). The coefficient estimates are obtained from Regression (9), replacing the $VolReg_t$ variables with $ModVolReg_t$. Unlike $VolReg_t$ (for which volatility after-effect theory predicts a negative coefficient), volatility after-effect theory does not predict a significant coefficient for $ModVolReg_t$. Columns report different values of the threshold that defines very high and very low volatility states (volatility that is greater than x standard deviation from the mean). Rows report different values of the threshold that defines the neutral volatility state (volatility that is within y standard deviations from the mean). t-statistics (in parenthesis) are calculated with heteroskedasticity-robust standard errors. ***, **, and * denote coefficients significant at the 1%, 5%, and 10% level respectively.

	x			
y	1.00	1.25	1.50	1.75
1.00	0.033	−0.039	0.160	−0.369
	(0.09)	(−0.10)	(0.32)	(−0.53)
1.25	.	0.240	0.411	−0.216
		(0.91)	(1.18)	(−0.46)
1.50	.	.	0.397	−0.191
			(1.38)	(−0.52)
1.75	.	.	.	−0.184
				(−0.60)

[27] See, e.g., Malmendier and Nagel (2011, 2016), Choi and Mertens (2019), Greenwood and Shleifer (2014), Barberis et al. (2015), and Frydman and Nave (2017).

[28] The average absolute difference between the log realized volatility in the neutral state and the log realized volatility in the very high or very low states is about the same for $ModVolReg_t$ regimes as it is for the $VolReg_t$ regimes so our statistical power is similar in both tests.

Table 5
Test of the Gambler's Fallacy hypothesis
This table reports coefficient estimates for the variable $NonVolReg_t$, which is non-zero if the volatility stays in the very high or very low state (in contrast to the $VolReg_t$ variables, which are non-zero after transitioning to the neutral state). The coefficient estimates are obtained from Regression (9), replacing the $VolReg_t$ variables with $NonVolReg_t$. The Gambler's Fallacy hypothesis predicts significant negative coefficients for $NonVolReg_t$. Columns report different length in days of the stimulus window. Rows report different values of the threshold that defines very high and very low volatility states (volatility that is greater than x standard deviation from the mean). t-statistics (in parenthesis) are calculated with heteroskedasticity-robust standard errors. ***, **, and * denote coefficients significant at the 1%, 5%, and 10% level respectively.

	Length of stimulus window		
x	2 days	3 days	4 days
1.00	−0.315*	−0.295	−0.412
	(−1.94)	(−1.47)	(−1.59)
1.25	−0.117	−0.050	−0.035
	(−0.53)	(−0.17)	(−0.09)
1.50	−0.289	−0.126	−0.303
	(−0.98)	(−0.29)	(−0.48)
1.75	−0.228	−0.254	−0.136
	(−0.52)	(−0.37)	(−0.13)

Next, we consider adaptive expectations about volatility levels, anchoring, and the gambler's fallacy. In the Internet Appendix A.6, we formally show that our results are strengthened (the estimated volatility after-effect is stronger) when assuming that agents have adaptive expectations about volatility levels rather than rational expectations. This is because the two phenomena (adaptive expectations about volatility levels versus volatility after-effect) work as antagonistic forces, distorting the VIX in opposite directions.

The same is true regarding the anchoring bias—making insufficient adjustments from the previous volatility level, which acts as a kind of "anchor" or reference point. By causing perceived volatility to be biased upward (downward) in the aftermath of a prolonged state of high (low) volatility, the anchoring bias distorts the VIX in the opposite direction to the after-effect. Therefore, by not incorporating the anchoring bias into our structural model, we *underestimate* the true extent of the VIX distortions that are caused by the after-effect.

Finally, a well-documented behavioural bias that should be accounted for is the gambler's fallacy.[29] Under the gambler's fallacy, expected volatility decreases (increases) toward the end of a prolonged state of high (low) volatility due to the belief that the run must end. This is analogous to believing "heads" is more likely on a coin toss after a run of "tails". If this bias is present in our data, VIX should systematically change during a period of persistent extreme volatility and revert some of this change after a transition occurs. To test for this possibility, we define a new variable ($NonVolReg_t$) that, in contrast to the $VolReg_t$ variables, takes non-zero values when realized volatility stays in the same very high or very low state *without* transitioning to the neutral state. We vary the duration of the stimulus window from two to four days. For a four-day window, $NonVolReg_t$ is defined as:

$$NonVolReg_t = \begin{cases} +1 \ if \ \{LnRV_{t-3},\ LnRV_{t-2},\ LnRV_{t-1}, LnRV_t\} = \{VH,\ VH, VH, VH\} \\ -1 \ if \ \{LnRV_{t-3},\ LnRV_{t-2},\ LnRV_{t-1}, LnRV_t\} = \{VL,\ VL, VL, VL\} \\ 0 \ otherwise \end{cases} . \tag{13}$$

Table 5 reports the coefficients of $NonVolReg_t$ when it replaces the $VolReg_t$ variables in Regression (9). If the gambler's fallacy bias distorts VIX in our data, the coefficient of $NonVolReg_t$ would be significantly negative. This is not the case in our data.

A further alternative explanation that we consider is neglected risks and the availability heuristic. One prominent explanation of financial crises, which is supported by strong empirical evidence, is based on the idea of neglected risks and the availability heuristic (Gennaioli et al., 2012; Goetzmann et al., 2016). One key insight from that literature is that a prolonged period with no crashes leads people to neglect crash risk because crash events that occur in the distant past are hard to recall from memory. Applied to a regime-shifting setting, such a belief pattern suggests that investors underestimate the possibility of the low volatility state after prolonged exposure to high volatility (i.e., toward the end of a high volatility episode), and the transition to the neutral volatility state reminds investors of the existence of the low volatility state. Similarly, the transition to the neutral volatility state after prolonged exposure to low volatility would remind investors of the existence of the high volatility state. This "neglected-risks hypothesis" is consistent with the aforementioned literature on neglected risks and the regime-switching learning model of Veronesi (1999).

However, the VIX distortions predicted by the neglected-risks hypothesis do not match those observed in our data in several respects. First, our finding that the VIX distortions increase with stimulus strength in the episodes of extreme volatility as predicted by volatility after-effect theory, is *not* predicted by the neglected-risks hypothesis.[30] Further, two direct implications of the neglected-risks hypothesis are not supported in the data. The first implication is that VIX is distorted upward toward the end of high volatility episodes (when the agent neglects the possibility of the low volatility state) and downward toward the end of low volatility episodes (when the agent neglects the possibility of the high volatility state). This is at odds with the data: the results show there is no drift one way or another during high or low volatility episodes. The second implication is that VIX is *not* distorted during the transition phases. The distortions only occur after a prolonged episode of high or low volatility, as a consequence of neglecting the opposite volatility state. The distortions should vanish upon entering the transition phase because the transition reminds the agent about the neglected state. That is, the temporal pattern of the VIX distortions predicted by the neglected-risks hypothesis does not match those observed in the data: we find that the distortions occur during the transition phases but not during the high or low volatility episodes.

2.6. Robustness tests

We conduct many robustness tests, which we report in the Internet Appendix A.3. For instance, we augment our regressions to include proxies for changes in the variance risk premium, overall market activity, overall market liquidity, microstructure effects, changes in intermediary constraints, as well as several other potential confounds (e.g., variables that measure the slope of the implied volatility curve such as "contango"). Our results are robust to the addition of all these extra control variables; the key coefficient estimates barely change.

In other robustness tests, we check that our findings are robust to using alternative specifications for the regression (e.g., using VIX and realized volatility in levels rather than in logs). We re-

[29] See, e.g., Tversky and Kahneman (1971, 1974), Gilovich et al. (1985), and Chen et al. (2016).

[30] It is the fact that the agent has not experienced a given state for a while that generates the neglect of that state; however, the volatility level that is experienced is not a determinant of such neglect.

calculate both the standard deviations of our regressions (replacing robust standard deviations with Newey-West standard deviations) and the volatility states (very high, very low, neutral) used to construct our main regime variable (e.g., changing the window used to compute the volatility of realized volatility, classifying the volatility states using percentiles rather than fractions of standard deviations of the realized volatility distribution, etc.). We also find that outliers do not drive the results (winsorizing the data strengthens the results).

3. Laboratory experiments

3.1. Experimental design

As explained in the Introduction, we separately measure the after-effect post-low volatility and post-high volatility in the experimental setting proposed by Payzan-LeNestour et al. (2016), both in the original condition ($N = 56$), in which VL was defined as 0.02, and VH, 0.45, and the field condition ($N = 132$), in which VL and VH are set to more realistic levels (VH = 0.25 in 1 session, 0.3 in 3 sessions, 0.4 in 1 session; VL = 0.04 in 4 sessions, 0.07 in 1 session). The field condition pools together existing data from a companion study (Payzan-LeNestour et al., 2021b) and data from one extra experimental session specifically run for this study ($N = 31$).

Each experimental session is performed in accordance with relevant guidelines and regulations. It is approved by the ethics panel at the [institution]. Written consent is obtained from all participants. The participants are undergraduate students from [institution] recruited through Orsee[31] (mean age 21.6 years, median 20; 42% male).

Each participant performs a total of 28 trials, 20 experimental, 3 diversion and 5 control trials, all randomly interleaved. In each experimental trial, the participant undergoes a 50 s "adaptation period" that involves passively viewing a Brownian motion depicting trajectories of a financial asset displayed as a moving line-plot that is dynamically re-drawn on screen from right to left. A small dot signals the start of the motion, which leaves a line as a visible trace of the dot's vertical movement over time. In half the experimental trials, the asset volatility in the adaptation period is VH. In the other half it is VL. During the test period, which immediately follows each adaptation period, the participant views a medium (0.1) volatility asset for 20 s and rates its perceived volatility on a five-point scale using the mouse pointer to click the relevant button. The timeline of an experimental trial is provided in Fig. 5, and a demo of the task can be found online.[32]

The diversion trials consist of medium volatility followed by either VL or VH in the test stimulus, and the control trials consist of a medium-medium transition. The mean value of the asset is varied across trials; the values used are 0, ±0.1, ±0.2, ±0.3, ±0.4 (in randomized order).

Each experimental session is performed in a darkened room on 21″ HP L2105tm computer monitors with a resolution of 1920×1080 pixels, with a frame rate of 60 Hz. The experimental programs are web applications custom coded in PHP.[33]

Upon arrival at the lab, the participants are asked to read the online task instructions and are encouraged to ask any clarifying questions to the experimenter. To avert well-known decision biases related to the use of rating scales (e.g., participants using the stimulus seen in the adaptation phase as a reference point to rate the test stimulus), the task instructions begin by pinning down the meaning of the volatility scale used in the task via exemplar stimuli that define the two extreme points on the scale (1: VL; 5: VH) and the middle point (3: medium). Specifically, we show a 1 min video of a VH / VL / medium volatility asset to explain what we mean by "very high volatility" / "very low volatility" / "medium or neutral volatility."

Evidence suggests that participants use the examples seen during the task instructions as reference points for their ratings, as intended. They indeed correctly rate the test stimuli (VH = 5; VL = 1) in the large majority (more than 85%) of the diversion trials.

The participants are provided with high incentives to fully engage in the task. Specifically, they are told in the task instructions that the experimenters will monitor their focus and gaze on the computer monitor throughout the task using a webcam installed on top of the computer (visible to the participant). They are further told that they will receive a fixed payment of $50 if they track all the stimuli displayed on screen; otherwise they will only receive a $5 show-up reward (which is to be provided to all participants irrespective of their behaviour during the experiment, as per the lab rules).[34]

Just before performing the task, the participants are re-explained the distinctive nature of the task. The need to balance high pace of reply while never replying randomly is particularly emphasized by the experimenter. Participants are warned that after a few missed trials (which occur when the participant fails to reply within the imparted time), the task will stop automatically. The participants are also reminded that their behaviour will be recorded by a webcam throughout the task and that they must keep their focus and gaze on the monitor during the entire experiment to get the $50 payment. Each experimental session takes approximately 90 mins overall (the task itself takes 45 min). Based on two real-time video inspectors, all participants comply with our experimental requirements and are accordingly paid $50.

Exclusion criterion: One may find it advisable to discard the data from four participants (one in the original condition, three in the field condition) who reply inappropriately in the diversion trials (VL rated three or above or VH rated three or below). The conclusions of all the tests reported below hold whether these participants are included in the analysis or not.

3.2. Main findings from the laboratory

For each participant, the after-effect post-low is defined as the difference between the mean reported volatility post-low (the reported volatility averaged across the trials in which the participant is exposed to VL in the adaptation phase) and the mean reported volatility in the control trials. Likewise, the after-effect post-high is quantified via the difference between the mean reported volatility in the control trials and the mean reported volatility post-high (the reported volatility averaged across the trials in which the participant is exposed to VH in the adaptation phase).

Table 6 reports descriptive statistics for the after-effect on a participant level (see also Figures A.5 to A.8 in Section A.8 of the Internet Appendix). Across participants, there are 47 missed trials, leaving 4701 trials for the analysis. The after-effect post-low is normally distributed (when inspected by histogram, and we cannot reject the null hypothesis in a Shapiro-Wilks test). The after-effect post-high is not normally distributed in the field condition. The tests reported below use the raw data for the after-effect, and we check that the conclusions of all tests are unchanged after trans-

[31] http://www.orsee.org/.

[32] https://player.vimeo.com/video/124360127?title=0amp;byline=0&portrait=0&color=c9ff23.

[33] The code to program the experimental task, as well as the code to replicate the analyses reported in the paper, and the experimental data used for the study, can be downloaded athttps://www.dropbox.com/sh/iqe6fs8mvtkm95h/AAB9MCLhFXWxUccaHIhS-qXSa?dl=0

[34] The task instructions can be found at http://instructionsna.weebly.com/.

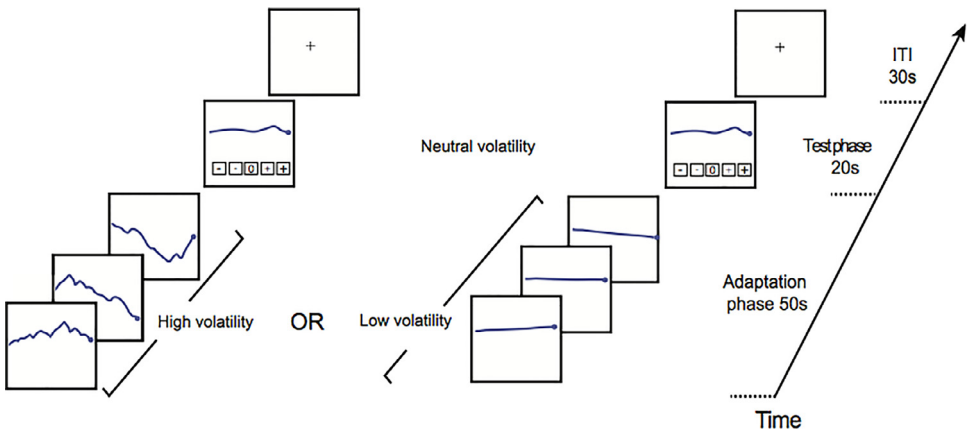

Fig. 5. Experimental design used in the laboratory experiment. Adaptation phase: Brownian motion with volatility level set at either VH (left) or VL (right) is displayed for 50 s. Test phase: participants report their perceived volatility of a medium (0.1) volatility Brownian motion using the mouse pointer. ITI: 30 s Inter Trial Interval.

Table 6
Descriptive statistics for the after-effect post-low and post-high volatility.

	Condition	N	Mean	StdDev	Median	Min	Max	Skew
Post-low	Original	56	0.34	0.50	0.25	−0.62	1.62	0.43
	Field	132	0.18	0.38	0.19	−1.00	1.30	−0.04
Post-high	Original	56	0.44	0.45	0.38	−0.50	1.62	0.81
	Field	132	0.36	0.45	0.30	−0.83	2.88	1.33

Table 7
One-sample two-tailed t-tests comparing the mean after-effect post-low to a population mean of 0.
The test is estimated for the original and field conditions separately. H_0: the after-effect post-low is null. Power to detect a medium ($D = 0.5$) effect size for a 5% alpha level is 96% for the original condition and 100% for the field condition.

Condition	Original	Field
Mean	0.34	0.18
t-statistic	5.04	5.34
p-value	0.000	0.000
P5	0.20	0.11
P95	0.47	0.24
Degrees of freedom	55	131
Cohen's D	0.67	0.46

Table 8
Frequency table for a one-tailed Fisher's exact test comparing the prevalence of the after-effect across conditions
H_0: the proportion of participants with a positive after-effect is the same or higher in the field condition. The test is run for the after-effect post-high and post-low separately. Post-high: odds ratio = 1.612, $p = 0.185$; post-low: odds ratio = 1.946, $p = 0.037$. Power to detect a medium effect size for a 5% alpha level is 100%. For reference, one may use 1.22, 1.86, and 3.00 to define small, medium, and large odds ratios in this test, following Oliver and Bell (2013).

		Post-high		Post-low	
		After-effect ≤ 0	After-effect > 0	After-effect ≤ 0	After-effect > 0
Condition	Original	8	48	15	41
	Field	28	104	55	77

forming the after-effect to reduce skew (we use a Box-Cox transformation with $\lambda = 3$).

We run one-sample two-tailed t-tests to assess the magnitude of the after-effect and find a significant after-effect, both post-low and post-high, in each condition. Table 7 reports the findings for the after-effect post-low.[35]

Table 9
Two-sample one-tailed *t*-test comparing the after-effect post-low across conditions
H_0: the after-effect post-low is higher or equal in the field condition than in the original condition. For this test, power to detect a medium effect for a 5% alpha level is 93%. We also report the minimum effect size (Cohen's D) needed to find a significant effect with pre-specified power and alpha, given our sample size ("Sensitivity analysis of power").

Mean Original	0.34
Mean Field	0.18
Mean difference	−0.16
t-statistic	−2.39
p-value	0.009
P5	-Infinity
P95	−0.05
Degrees of freedom	186
Cohen's D	−0.36

Sensitivity analysis of power:

	Alpha = 5%	Alpha = 10%
Power = 80%	0.449	0.398
Power = 90%	0.520	0.468

The after-effect post-low appears to be weaker in the field condition relative to that in the original condition, in terms of prevalence (Table 8) and magnitude (see Table 9; the results of a non-parametric Wilcoxon Rank Sum test lead to the same conclusion: $W = 3024$, $p = 0.024$).

Consistent with these results, we find that in a mixed-effects model estimating the after-effect post-low, a dummy regressor to compare the original condition (the reference) to the field condition is significantly negative (Table 10, Regression (1)). We further

[35] The corresponding table for the after-effect post-high is omitted here in the interest of space.

Table 10
Regression table for mixed-effects models estimating the after-effect
"Condition (Field)": dummy to compare the original condition (reference) to the field condition. "Volatility": VL (0.02, 0.04, or 0.07) for the after-effect post-low, VH (0.25, 0.30, 0.40, 0.45) for the after-effect post-high. The models include by-participant intercepts. After-effect is z-scored in all regressions and volatility is z-scored in regressions (2) and (4). ***, **, and * denote coefficients significant at $p < 0.001$, $p < 0.01$, and $p < 0.05$ respectively.

	Post-low		Post-high	
	(1)	(2)	(3)	(4)
Intercept	0.16 (0.08) *	0.00 (0.04)	0.09 (0.08)	0.00 (0.05)
Condition (Field)	−0.23 (0.10) **	–	−0.12 (0.10)	–
Volatility	–	−0.10 (0.04) **	–	0.07 (0.04)
R^2	0.28	.27	.31	.31
Observations	1687	1687	1691	1691

Table 11
Power for the main variables of interest in the regressions reported in Table 10
Power to detect a small and medium effect sizes for a 5% alpha level is presented as (small, medium). Power is estimated using 1000 simulated datasets with a pre-specified effect size (small: $\beta = -0.20$) for the condition/volatility variables.

	Post-low		Post-high	
	(1)	(2)	(3)	(4)
Condition (Field)	(27%, 90%)	–	(27%, 90%)	–
Volatility	–	(76%, 100%)	–	(76%, 100%)

find that VL value is a predictor of the magnitude of the after-effect post-low (Table 10, Regression (2)). In contrast, we do not find evidence that the magnitude of the after-effect post-high is reduced in the field condition (Table 10, Regression (3)) or sensitive to VH value (Table 10, Regression (4)).

In summary, these findings suggest that the after-effect post-low is present in the field condition, though it is weaker than in the original condition—perhaps not surprisingly given the extremely low value of VL used in the original condition (recall that VL = 0.02) and given that the effect appears to be sensitive to the VL value. The after-effect post-high appears to be more robust in our laboratory data, similar to the field evidence. However, we caution against concluding that its magnitude is not sensitive to the level of VH; it could be sensitive to it, but the effect would be too small for us to be able to detect it in our tests due to insufficient sample size Table 11).[36]

In the final stage of analysis, we test whether the after-effect is asymmetric. The results of a paired one-tailed *t*-test comparing the mean after-effect post-high to the mean after-effect post-low on a participant level suggest that the magnitude of the after-effect is stronger post-high (Table 12). The results of a McNemar test suggest that the prevalence of the after-effect is higher post-high (Table 13).

We conclude from the collection of findings that the asymmetry observed in the field is unlikely to come from the fact that we do not have a sufficiently large number of VL regimes for equity volatility. Rather, the evidence suggests that the after-effect is inherently stronger post-high, though it is nonetheless significant post-low (both in the whole sample and the sample restricted to the field condition), suggesting that both sides of the after-effect might constitute an exploitable arbitrage opportunity (more in Discussion below).

[36] Table 6 shows that if the effect for the main variables of interest in the regressions (condition and volatility) turned out to be small, then the regressions reported in Table 6 would be underpowered. This means that the absence of evidence that the condition and volatility variables are significant predictors of the after-effect post-high in Regressions ((3) and (4) should *not* be interpreted as evidence of absence of an effect. The effect could be too small for us to identify it here.

Table 12
Paired one-tailed *t*-test comparing the mean after-effect post-high to the mean after-effect post-low at the participant level
The test is restricted to the participants in the field condition for whom the after-effect is not negative both sides. This leaves 123 participants out of 132. Power to detect a small and medium effect sizes for a 5% alpha level is 71% and 100% respectively. We also report the minimum effect size needed to find a significant effect with pre-specified power and alpha, given our sample size for this test ("sensitivity analysis of power").

Mean post-high	0.40
Mean post-low	0.20
Mean difference	0.20
t-statistic	3.36
p-value	0.000
P5	0.10
P95	Infinity
Degrees of freedom	122
Cohen's D	0.30

Sensitivity analysis of power:

	Alpha = 5%	Alpha = 10%
Power = 80%	0.225	0.192
Power = 90%	0.265	0.231

4. Discussion

We find that the after-effect consistently affects volatility perception in the laboratory and the field. The after-effect following high volatility appears to distort asset prices in one of the most actively traded markets in the world, which points to its pervasiveness. We further show that the after-effect following low volatility is present in the laboratory (not statistically significant in field data), although not as strong as after-effects following high volatility.

The practical relevance of these findings is that they imply traders make systematic errors in assessing and responding to volatility following periods of very low or very high volatility levels. A follow-up study shows that the after-effect post-high constitutes an exploitable arbitrage opportunity, especially in the last ten years (Payzan-LeNestour et al., 2022). Our laboratory findings point to the after-effect post-low being likewise exploitable in markets where volatility levels regularly fall to levels around 4% and below.

Table 13
Frequency table used for a MacNemar test used to assess if the after-effect is asymmetric
H_0: the probability of the after-effect post-low being negative and the after-effect post-high being positive is equal or smaller than the probability of the after-effect post-low being positive and the after-effect post-high being negative. $\chi^2(1) = 12.375$, $p < 0.001$, OR = 1.714. Power to detect a small and medium effect sizes for a 5% alpha level is 12% and 86% respectively. The table also reports the minimum odds ratio needed to detect an effect for given power and alpha, given our sample size for this test. The test uses the data from the two conditions pooled together. The same test run on the field condition leads to the same conclusion ($\chi^2(1) = 10.4$, $p = 0.001$, OR = 2.421) but it is underpowered (power to detect a medium effect size for a 5% alpha level is 74%), so not reported here. *Note*: For reference, one may use 1.22, 1.86, and 3.00 to define small, medium, and large odds ratios in this test, following Oliver and Bell (2013).

	Post-high ≤ 0	Post-high > 0
Post-low ≤ 0	9	61
Post-low > 0	27	91

Sensitivity analysis of power:

	Alpha = 5%	Alpha = 10%
Power = 80%	1.78	1.65
Power = 90%	1.93	1.80

A natural question raised by the current findings is whether the price distortions generated by the after-effect will decline subsequent to dissemination of this research, as is the case with other stock market "anomalies" (e.g., McLean and Pontiff, 2016). Market participants may be able to consciously correct their decisions to account for the after-effect once they have been made aware of it. Addressing this question will require revisiting the empirical analysis in the future once additional data have been generated.

Another question raised by our findings is whether we should expect the price distortions caused by the after-effect to diminish through time with the increasing automation of trading. Based on practitioner reports of the trading process and recent evidence on manipulation of the VIX (Griffin and Shams, 2018), we believe the answer is not necessarily. Although the process of trade execution and market maker quoting is largely automated, the trading decisions that ultimately determine price levels are still driven by humans. This is true of many markets including the S&P500 options market that underpins the VIX. Automation has become widespread in the *mechanical* aspects of trading, such as market making algorithms that quote prices based on the prices of other assets and the market maker's inventory, trade execution algorithms that take an order (from a human) and optimally slice it up and execute it strategically to minimize costs, and arbitrage algorithms that seek out and exploit relative mispricing. But in contrast, humans continue to control tasks that involve *judgement*, such as key trading decisions, forecasting future prices including anticipating the actions of other traders, and gauging whether price levels are correct. For example, in the S&P500 options market, while algorithms ensure that options with different strikes and maturities are correctly priced relative to one another, humans often trade based on their beliefs about future volatility and judgements about whether the level of implied volatility is too high or too low. These judgement tasks, which are susceptible to distortions from perceptual errors such as the after-effect, are ultimately what determine price levels.[37]

The methodology used in the study is best described as a "loop process" whereby we start in the lab, then turn to the field to check the external validity of the lab results, and ultimately return to the lab to better understand features of the field evidence. This approach illustrates the complementarity of laboratory and field data. The original lab experiment provides us with a solid prior regarding the existence and direction of the effect investigated in the field. However, the laboratory findings present inherent limits related to external validity.[38] Moreover, by nature they cannot speak to the pervasiveness of the effect. The field data are crucial in both regards—to strengthen the evidence of the existence of the effect and to study its pervasiveness. They also allow us to identify a feature of the effect that is not implied by after-effect theory per se, namely, its asymmetry.

However, the field data present their own set of limits as well, which relate to the fact that there are many confounding factors in the field. Although we show that the patterns of VIX distortions created by the after-effect run in the *opposite* direction to what one might consider the most standard deviations from rational expectations, namely anchoring effects, adaptive expectations about volatility levels, and neglected risks, one cannot rule out that other factors contribute to the VIX distortions that we document. The evidence gathered from field data *alone* is thus inherently limited. This explains why it is crucial that the field study starts with a strong prior hypothesis that is grounded both in theory and from gathering laboratory evidence in a first step.

Moreover, the field data are ambiguous regarding an asymmetry that we find in the effects (whether it is real or simply reflects limited power). By offering the advantage of controlled experimentation, the lab can be used to address this kind of question, as we propose in the final stage of this study. Lab and field data are thus highly complementary (List, 2007). We hope the current study will inspire more experimental finance studies conducted in the same vein.

Declaration of Competing Interest

Elise Payzan-LeNestour: I have nothing to disclose.
Lionnel Pradier: I have nothing to disclose.
Talis Putnins: I have nothing to disclose.

CRediT authorship contribution statement

Elise Payzan-LeNestour: Conceptualization, Methodology, Investigation, Writing – review & editing. **Lionnel Pradier:** Conceptualization, Methodology, Investigation. **Tālis J. Putniņš:** Conceptualization, Methodology, Investigation, Writing – review & editing.

Data availability

Link to code and data provided in the manuscript

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jbankfin.2022.106685.

References

Baele, L., Driessen, J., Ebert, S., Londono, J.M., Spalt, O., 2019. Cumulative prospect theory, option returns, and the variance premium. Rev. Financ. Stud. 32, 3667–3723.
Barberis, N., Shleifer, A., 2002. Style investing. J. Financ. Econ. 68, 161–199.
Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015. X-CAPM: an extrapolative capital asset pricing model. J. Financ. Econ. 115, 1–24.

[37] We thank a number of practitioners for providing insights about the trading process (James Doran, Shuo Song, David Rabinowitz, and Christian Daher). Their views consistently emphasized the important role of human decision making despite automation of parts of the trading process.

[38] For example, results obtained with undergraduate students do not necessarily apply to professional investors (although it seems they do at least in some cases; for example, Fréchette 2015). Another limitation of the experimental part of this study is that the task participants are not remunerated based on the accuracy of their replies. Therefore, like in self-report questionnaires, there is a hazard that participant replies do not faithfully represent their perception.

Barlow, H.B., Jameson, D., Hurvich, L.M., 1972. Dark and light adaptation: psychophysics. Handbook of Sensory Physiology (eds.). Springer-Verlag, New York.
Barlow, H.B., Hill, R.M., 1963. Evidence for a physiological explanation of the waterfall phenomenon and figural after-effects. Nature 200, 1345–1347.
Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. Rev. Financ. Stud. 22, 4463–4492.
Bollerslev, T., Gibson, M., Zhou, H., 2011. Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. J. Econom. 160, 235–245.
Borsboom, C., Janssen, D.J., Strucks, M., Zeisberger, S., 2021. History matters: how short-term price charts hurt investment performance. J. Bank. Finance, 106351 (forthcoming).
Burr, D., Ross, J., 2008. A visual sense of number. Curr. Biol. 18, 425–428.
Carr, P., Wu, L., 2006. A tale of two indices. J. Deriv. 13, 13–29.
Carr, P., Wu, L., 2009. Variance risk premiums. Rev. Financ. Stud. 22, 1311–1341.
Chen, D.L., Moskowitz, T.J., Shue, K., 2016. Decision-making under the gambler's fallacy: evidence from asylum judges, loan officers, and baseball umpires. Q. J. Econ. 131, 1181–1242.
Corsi, F., Mittnik, S., Pigorsch, C., Pigorsch, U., 2008. The volatility of realized volatility. Econom. Rev. 27, 46–78.
De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990. Noise trader risk in financial markets. J. Polit. Econ. 98, 703–738.
Fleming, J., Ostdiek, B., Whaley, R.E., 1995. Predicting stock market volatility: a new measure. J. Futur. Mark. 15, 265–302.
Froot, K.A., Dabora, E.M., 1999. How are stock prices affected by the location of trade? J. Financ. Econ. 53, 189–216.
Fréchette, G., 2015. Laboratory experiments: professionals versus students. Handbook of Experimental Economic Methodology. Oxford University Press.
Gennaioli, N., Shleifer, A., 2018. A Crisis of Beliefs: Investor Psychology and Financial Fragility. Princeton University Press, Princeton, NJ.
Gennaioli, N., Shleifer, A., Vishny, R., 2012. Neglected risks, financial innovation, and financial fragility. J. Financ. Econ. 104, 452–468.
Gennaioli, N., Shleifer, A., Vishny, R., 2015. Neglected risks: the psychology of financial crises. Am. Econ. Rev. 105, 310–314.
Gilovich, T., Vallone, R., Tversky, A., 1985. The hot hand in basketball: on the misperception of random sequences. Cogn. Psychol. 17, 295–314.
Glimcher, P.W., 2014. Understanding the hows and whys of decision-making: from expected utility to divisive normalization. Cold Spring Harb. Symp. Quant. Biol. 169–176 Cold Spring Harbor Laboratory Press.
W.N. Goetzmann, D. Kim, and R.J. Shiller, 2016, Crash beliefs from investor surveys, *Working paper*.
Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. Rev. Financ. Stud. 27, 714–746.
Griffin, J.M., Shams, A., 2018. Manipulation in the VIX? Rev. Financ. Stud. 31, 1377–1417.
Hartzmark, S.M., Shue, K., 2018. A tough act to follow: contrast effects in financial markets. J. Finance 73, 1567–1613.
Hershenson, M.B., 1989. Duration, time constant, and decay of the linear motion aftereffect as a function of inspection duration. Percept. Psychophys. 45, 251–257.
Holzmeister, F., Huber, J., Kirchler, M., Lindner, F., Weitzel, U., Zeisberger, S., 2020. What drives risk perception? A global survey with financial professionals and laypeople. Manag. Sci. 66 (9), 3977–4002.
Huber, C., Huber, J., Kirchler, M., 2021. Market shocks and professionals' investment behavior – evidence from the COVID-19 crash. J. Bank. Financ. 133.
Huber, J., Palan, S., Zeisberger, S., 2019. Does investor risk perception drive asset prices in markets? Experimental evidence. J. Bank. Financ. 108.
Hurvich, L.M., Jameson, D., 1957. An opponent-process theory of color vision. Psychol. Rev. 64, 384–404.
Jiang, G.J., Tian, Y.S., 2007. Extracting model-free volatility from option prices: an examination of the VIX index. J. Deriv. 14, 35–60.
Khaw, M.W., Glimcher, P.W., Louie, K., 2017. Normalized value coding explains dynamic adaptation in the human valuation process. Proc. Natl. Acad. Sci. 114, 12696–12701.
Klos, A., Weber, E.U., Weber, M., 2005. Investment decisions and time horizon: risk perception and risk behavior in repeated gambles. Manag. Sci. 51 (12), 1777–1790.
Lamont, O.A., Thaler, R.H., 2003. Can the market add and subtract? Mispricing in tech stock carve-outs. J. Polit. Econ. 111, 227–268.
C. Laudenbach, S. Siegel, M. Strucks, and S. Zeisberger, 2022, Why do people (not) invest? The role of return and risk expectations, *SSRN working paper*.
Lee, C.M.C., Shleifer, A., Thaler, R.H., 1991. Investor sentiment and the closed-end fund puzzle. J. Finance 46, 75–109.
Leopold, D.A., Rhodes, G., Müller, K.M., Jeffery, L., 2005. The dynamics of visual adaptation to faces. R. Soc. Proc. B 272, 897–904.
List John A., 2007. "Field Experiments: A Bridge between Lab and Naturally Occurring Data," The B.E. Journal of Economic Analysis & Policy, De Gruyter, vol. 5(2), pages 1-47, April.
Laughlin, Simon. "A Simple Coding Procedure Enhances a Neuron's Information Capacity" Zeitschrift für Naturforschung C, vol. 36, no. 9–10, 1981, pp. 910–912. doi:10.1515/znc-1981-9-1040.
Magnussen, S., Johnsen, T., 1986. Temporal aspects of spatial adaptation. A study of the tilt aftereffect. Vis. Res. 26, 661–672.
Malmendier, U., Nagel, S., 2011. Depression babies: do macroeconomic experiences affect risk-taking? Q. J. Econ. 126, 373–416.
Mata, R., Frey, R., Richter, D., Schupp, J., Hertwig, R., 2018. Risk preference: a view from psychology. J. Econ. Perspect. 32 (2), 155–172.
McLean, R.D., Pontiff, J., 2016. Does academic research destroy stock return predictability? J. Finance 71, 5–32.
Merton, R.C., 1980. On estimating the expected return on the market: an exploratory investigation. J. Financ. Econ. 8, 323–361.
Mitchell, M., Pulvino, T., Stafford, E., 2002. Limited arbitrage in equity markets. J. Finance 57, 551–584.
Payzan-LeNestour, E., Balleine, B., Berrada, T., Pearson, J., 2016. Variance after-effects bias risk perception in humans. Curr. Biol. 26, 1500–1504.
Payzan-LeNestour, E., Pradier, L., Putnins, T., 2022. Harnessing neuroscientific in--sights to generate alpha. Financ. Anal. J. 78 (2), 79–95 *R&R*.
Payzan-LeNestour, E., Pradier, L., Doran, J., Nave, G., Balleine, B., 2021b. Impact of ambient sound on risk perception in humans: neuroeconomic investigations. Sci. Rep. 11, 5392.
Payzan-LeNestour, E., Woodford, M., 2022. Outlier blindness: a neurobiological foundation for neglect of financial risk. J. Financ. Econ. 143, 1316–1343.
Shleifer, A., Vishny, R.W., 1997. The limits of arbitrage. J. Finance 52, 35–55.
Sitkin, S.B., Weingart, L.R., 1995. Determinants of risky decision making behavior: a test of the mediating role of risk perceptions and risk propensity. Acad. Manag. J. 38, 1573–1592.
Slovic, P., Bazerman, M., Messick, D.,, Tenbrunsel, A., Wade-Benzoni, K., 1997. Trust, emotion, sex, politics, and science: surveying the risk-assessment battlefield. In: Psychological Perspectives to Environmental and Ethical Issues in Management. Jossey-Bass, San Francisco, CA, pp. 277–313 (eds).
Tversky, A., Kahneman, D., 1971. Belief in the law of small numbers. Psychol. Bull. 76, 105–110.
Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. Science 185, 1124–1131.
Veronesi, P., 1999. Stock market overreaction to bad news in good times: a rational expectations equilibrium model. Rev. Financ. Stud. 12, 975–1007.
Unser, M., 2000. Lower partial moments as measures of perceived risk: an experimental study. J. Econ. Psychol. 21 (3), 253–280.
Weber, E.U., Shanteau, J., Mellers, B.A., Schum, D., 1998. Who's afraid of a little risk? New evidence for general risk aversion. Decision Research from Bayesian Approaches to Normative Systems: Reflections on the Contributions of Ward Edwards. Kluwer Academic Press, Norwell, MA (eds).
Weber, E.U., Blais, A.R., Betz, N.E., 2002. A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors. J. Behav. Decis. Mak. 15 (4), 263–290.
Weber, E.U., Hsee, C., 1998. Cross-cultural differences in risk perception, but cross–cultural similarities in attitudes towards perceived risk. Manag. Sci. 44 (9), 1205–1217.
Weber, E.U., Milliman, R.A., 1997. Perceived risk attitudes: relating risk perception to risky choice. Manag. Sci. 43 (2), 123–144.
Weber, M., Weber, E.U., Nosic, A., 2013. Who takes risks when and why: determinants of changes in investor risk-taking. Rev. Financ. 17 (3), 847–883.
Zeisberger, S., 2022. Do people care about loss probabilities? J. Risk Uncertain..