The Journal of Behavioral Finance & Economics Volume 3, Issue 2, Fall 2013
**25-42** 
ISSN: 1551-9570

# Asset Pricing Theory in Light of Ellsberg Paradox

Peter B. Lerner[1]
*Scitech Solar Analytics*

**Abstract**

Behavioral patterns frequently contradict rational expectations. Yet, quantitative consequences of violation of rational expectations have been so far obscure. In 2008, Epstein and Schneider proposed a theory, which incorporates classical the Ellsberg paradox into a rigorous asset pricing model. In 2009, the author pointed out that Epstein-Schneider asset pricing theory might explain an extremely high sensitivity of market indexes to the changes in a risk free rate, though contribution of the risk free rate to the cost of capital of a high-tech (for instance, a typical Nasdaq firm) can be quite small. In this paper, I subject this circumstance to an empirical verification using intraday integrated volatility of the Nasdaq-100 index.

***JEL***: G02, G12, G14

[1] Scitech Solar Analytics, LLC, State College, PA 16802. The author thanks his discussant at the 18th Annual Multinational Finance Society Conference, Asli Ascioglu from Bryant University and Petko Kalev (University of South Australia), the chair of the session, for valuable suggestions and the participants of discussion at the 4th Annual Meeting of Academy of Behavioral Finance and Economics. All errors are my own.

**Introduction**

The Ellsberg paradox and its variants are classical manifestations of the incompatibility of market reasoning with a rational expectations paradigm. (Fox and Tversky, 1995) Namely, the human mind clearly assigns different value to hidden but certain information and "true" statistical uncertainty. The classic Ellsberg paradox is formulated as follows: consider four versions of the pebbles-in-the-urns game. The reward matrix is given by the table

| | |
|---|---|
| R | B |
| R V Y | R V B |

**Note**: All urns contain 30 red (R) balls and 60 balls, which can be black (B) or yellow (Y). Reward is assumed to be the same for all gambles.

The majority of the players prefer rewards located on the diagonal of the table in direct contradiction to the existence of a unique utility function. Obviously, when in doubt, a human observer tends to expect the worst outcome even if she assumed the opposite outcome in situation that is statistically equivalent but contains different distinction between "hidden" and "unavailable" information.

However, practical implications of this and other behavioral phenomena are hard to establish beyond proverbial "madness of the crowds." (McKay, 2011) In 2008, Epstein and Schneider (Epstein, Schneider, 2008, further ES) formulated a microstructure-inspired theory in which one could determine price volatility through a number of other market parameters such as asset volatility, risk free rate and dividend rate. The foundation of their theory is the idea that all market risk is not created equal but consists of two components. One component of risk is the real uncertainty in a future payment stream from an asset. Another is the limited possibility of a participant to distinguish a real event of economic significance from market noise.

To formulate an asset pricing model, which involves market noise from the outset, one has to evoke market microstructure concepts. The problem with all microstructure asset pricing theories is that they contain a number of unobservable parameters (O'Hara, 1995, Hasbrouck, 2007), which makes a comparison of microstructure theories with empirical data exceedingly difficult if not impossible. The Epstein-Schneider theory provides a bridge from microstructure models to the basic financial concepts of the time value of money and the risk-free rate. These concepts are imperative for microstructure-based valuation models. Because, as elsewhere in microeconomics (Mas-Colell, Whinston and Green, 1995), the time value of money is a consequence of risk avoidance of market agents, the solution to this problem needs to incorporate risk avoidance, or betting on the worst-case scenario.

The most striking practical consequence of the ES is the extreme sensitivity of market volatility of a non-dividend paying stock to a risk free rate. This consequence can be empirically tested using intraday integrated volatility of any popular stock index. I begin my study with

Nasdaq-100 because the valuation of the high-tech stocks comprising majority of the index are not expected to depend very much on the risk free rate since it contributes only a minuscule portion to the cost of capital of high tech firms and many of them have to raise capital from private sources only indirectly related to the US Treasury market.

The strategy of this paper is rather straightforward: regress (logarithmic) market volatility of a market index obtained from intraday quotes on a risk free (short) rate. However, assessment of integrated volatility from the intraday quotes in the presence of microstructure noise poses quite a few econometric problems (Andersen, 2003) some of which will be addressed in the main text of the paper and the Appendixes. The paper is structured as follows. In the first section, I recount a simplified version of ES sufficient for our purposes. In the second section, I study a sample of integrated daily volatility during 25 trading days of Nasdaq-100 from 2005. In the third section, I make a benchmark test of the hypothesis in the relatively quiet summer 2005, not selected by any drastic events in the money market. In the fourth section, I study the days surrounding the Standard-and-Poors downgrade of the US AAA credit rating on August 5, 2011. I discuss the paper's results in a Conclusion section.

## 1. Exposition of the Epstein-Schneider theory

I shall demonstrate a simplified version of the ES model in this section. Their argument runs as follows. Let us distinguish between two kinds of price uncertainty: noisy prices and ambiguity. With the first, Epstein and Schneider (ES, 2008) acknowledge our lack of understanding of the true value of an asset. With the second, they acknowledge our limited capacity to interpret price signals.

From the point of view of a Bayesian observer, these two sources of uncertainty can be unified by a proper choice of prior. However, if the observer is risk-avoidant, under the same expected payoff the agent prefers a terminal state with the least total uncertainty. If she cannot distinguish between true asset volatility and microstructure noise, the noisier state may win.

To quantify this distinction, ES consider a model with a dividend-paying asset. To simplify matters, I consider an asset that pays no dividends, such as a market index. Hence, I consider that a true price (return, in the case of stocks) follows a Kyle-type (Kyle, 1985) equality:

$$p_t = p_0 + u_t \tag{1}$$

where the asset price uncertainty $u_t \sim N(0, \sigma^2_u)$. Without limitations on the generality we can consider $p_0=0$ and denote the (anomalous) return of an asset as $q_t$.

In contrast to the Kyle model, where the equilibrium is always revealed in the end of each trading session, in the ES model learning is possible only to the extent of the observation error $e_t$. The assumption of the normal distribution of observation error $e_t \sim N(0, \sigma^2_s)$ is technical and was made by ES for analytical tractability. In their approach, $\sigma_s$ is a definite number that is not amenable to direct observation. The observed price $s_t=p_t+u_t$ is also normally distributed with the zero mean and variance $\sigma_u^2+\sigma_s^2$. Price volatility is assumed as being unknown but fixed.

True stock price is revealed in a series of auctions, in which market price is observed in the following manner. A market participant ascribes to the observed price *s* her "strength of conviction" variable with a binary distribution with the upper and lower values $\underline{\gamma}$ and $\bar{\gamma}$ between 0 and 1. For single auction, the price presumption for a risk avoidant agent is equal to the worst-case scenario for the current observation:

$$q_1 = m + \underline{\gamma} s + (\bar{\gamma} - \underline{\gamma}) \min(s,0) \qquad (2)$$

where *m* is a mean price. We denote the expectation of γ under a true probability distribution as γ*. In the efficient market the expected mean is zero: *E[m]=0*. Below we consider the limit of infinitely many auctions. The price process thus acquires the form:

$$q_{t+1} = Q_0 + Q_1 \gamma_t s_t \qquad (3)$$

Undetermined coefficients $Q_0$, $Q_1$ in Equation (3) are determined from the Euler-Lagrange equation (Mas-Colell, Whinston and Green, 1995):

$$q_t = \min_{\sigma_s^2 \in [\underline{\sigma}_s^2, \bar{\sigma}_s^2]} E_t \left[ \frac{1}{1+r} q_{t+1} \right] \qquad (4)$$

where *r* is a risk-free rate. The Euler-Lagrange equation, as always, is a consequence of the fact that immediate consumption is valued more than the consumption at a later time by the factor of risk avoidance $\beta = (1+r)^{-1}$.

The solution of Equation (4) provides

$$Q_0 = -\frac{(\bar{\gamma} - \underline{\gamma})\sigma_u}{r\sqrt{2\pi\gamma^*}} \qquad Q_1 = \frac{1}{r} \qquad (5)$$

In (5), γ* is the expectation of random variable γ under an unbiased distribution. The first term in Equation (5) can thus be interpreted as an anticipation of future ambiguous news and the second term as the response to the current ambiguous signal. The expression for the return volatility Σ, which we identify with the standard deviation of the returns,

$$r_{t+1} = q_{t+1} - (1+r) \cdot q_t$$

is equal to (ES, Section C.2):

$$\Sigma = \mathrm{var}[q_t] = \left( \frac{\sigma_u^2}{r^2} \right) \left( 1 - \bar{\gamma} - \underline{\gamma} - \frac{1+\beta^2}{2 \cdot \gamma^*} \left( \bar{\gamma}^2 + \underline{\gamma}^2 - \frac{1}{\pi} (\bar{\gamma} - \underline{\gamma})^2 \right) \right) \qquad (6)$$

Equation (6) reflects the difference of the situation with the Bayesian (risk-neutral) learning. Price volatility monotonically increases with asset volatility as well as with the uncertainty-of-beliefs parameters $\bar{\gamma}$, $\underline{\gamma}$ and $\xi = (\bar{\gamma} - \underline{\gamma})$ .

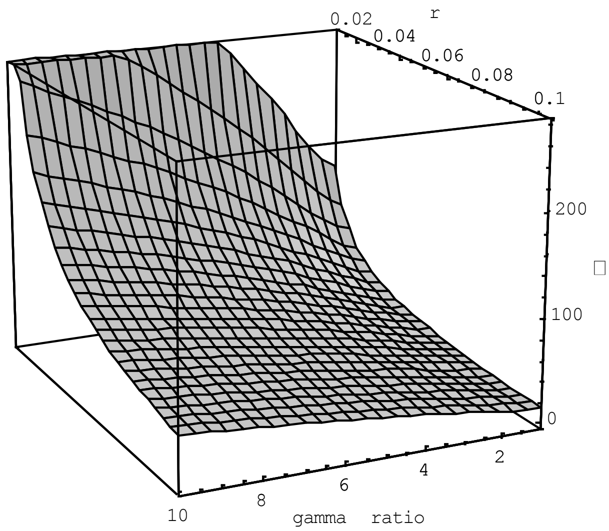


**Fig.1.** Volatility surface of the stock s=$\sqrt{\Sigma}$ (in units of asset volatility) as a function of risk-free rate r and ambiguity expressed as gamma ratio $\gamma_r = \bar{\gamma} / \underline{\gamma}$ (Equation 6).

This fact stands at variance with the Bayesian case, where price volatility increases with asset volatility but diminishes with learning uncertainty (Lerner, 2009). One difference between ES results for volatility and Bayesian theory is that in ES all agents possess essentially the same information, even if this information is ambiguous. According to standard versions of the informed trading model, the agents sharply differ with respect to privileged access to information. In effect, unlike Kyle and Glosten-Milgrom (1985) models, the participants in the ES game do not learn from the leader.

As one can observe from Fig. 1, a feature of the ES model is that for a no-dividend (or small dividend), stock price is very elastic with respect to the risk-free rate. A related feature is the extreme amplification of the asset volatility in the ES model: price volatility for quite reasonable values of parameters can be two orders of magnitudes higher than the asset volatility! In the next section, I use the specially filtered intraday returns on Nasdaq index as the proxy for the stock market and fed funds rate as a proxy for the risk-free rate. I calculate integrated daily volatility as annualized standard deviation of filtered intraday returns.

## 2. Determination of the volatility of the Nasdaq market

Volatility is one of the most ubiquitous concepts in finance yet the one is very difficult to measure empirically. In particular, volatility from intra-day observations can be highly unstable. (Fig. 2) The problem of choosing an optimum sampling frequency has been discussed and solved by Aït-Sahalia, Mykland and Zhang (SMZ, 2005, ZMS, 2005) and Bandi and Russell (2005, 2006) yielding comparable conclusions. I shall follow the approach of Bandi and Russell, which I find more instructive despite its reliance on lengthy calculations. A brief outline of the Bandi and Russell theory is provided in the Appendix A.

In this section I present my empirical results on the filtering with estimation frequency according to SMZ and Bandi-Russell. In Fig. 2, I present a plot of volatility of the Nasdaq-100 index during 25 inconspicuous trading days between 04/26/2005 and 05/31/2005 taken with 5-minute intervals (2175 data points) from Bloomberg©.

**Table 1 Descriptive statistics of the intraday returns on Nasdaq100 index (04/26/05-05/31/05). There is nothing particularly interesting in this sample except for the very pronounced kurtosis, i.e. the "fat tails" of the high frequency returns. Small skewness indicates upward motion of the index in the period of observation.**

| Mean | Standard Error | Median | Mode | Standard Deviation | Sample Variance | Kurtosis | Skewness |
|---|---|---|---|---|---|---|---|
| 2.81E-05 | 1.71E-05 | 0 | 0 | 0.000795 | 6.33E-07 | 25.84 | 1.314 |
| Range | Sum | Largest(1) | Smallest(1) | Confidence Level (95%) | | | |
| 0.0180 | 0.0611 | 0.010869 | -0.0071 | 3.35E-05 | | | |

The jagged curve in Fig. 2 depicts volatility calculated as a moving average of intraday data on any given day. The curves indicating 10-day, 30-day, 50-day and 100-day moving averages also provided by Bloomberg© are plotted on the same scale to illustrate the relationship between intraday and averaged daily volatility. In general, despite comparable levels of volatility there is no reciprocity between these graphs, and only 10-day moving averages demonstrate any semblance to the dynamics of the intraday volatility curves.

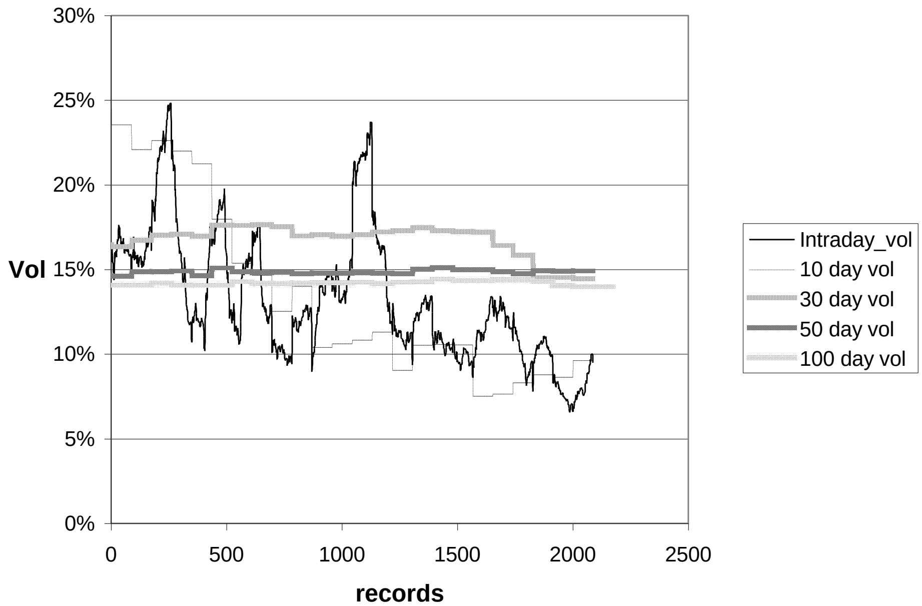


**Fig. 2.** Plot of Nasdaq-100 volatility taken in 5-minute intervals for 25 trading days from 04/26/2005 to 05/31/2005. Solid black curve is the intraday volatility. Fine black curve indicates 10-day moving average, shaded curves indicate 30-day (grey), 50-day (dark grey) and 100-day (light grey).

The Bandi-Russell algorithm suggests using optimal sampling frequency determined by the cubic equation (A.7 in the Appendix A). In the limit of a large number of observations, the solution of the cubic equation can be approximated by the following Equation (Equation (18) in Bandi and Russell (2006)):

$$c^* = \left(\frac{\hat{Q}}{\hat{\alpha}}\right)^{1/3} \tag{7}$$

where $\hat{Q}$ and $\hat{\alpha}$ are sample estimators of the parameters of $Q$ and $\alpha$. The formulas for the sample estimators are given by Equations (19-20) from Bandi and Russell, but one could have inferred them by applying Theorem 1 to the expressions of Theorem 2:

$$\hat{\alpha} = \left( \sum_{i=1}^{n} \sum_{j=1}^{M} \tilde{r}_{ji}^{2} \Big/ nM \right)^2 \tag{8}$$

$$\hat{Q} = M/3 \sum_{j=1}^{M} \tilde{r}_{ji}^{4}$$

The optimal sampling frequency computed by the Equation (8) (compare with Equation (9) in ZMS) is $K=c*M^{1/3}=5.99\approx 6$. An optimally filtered volatility is shown in Fig. 3.

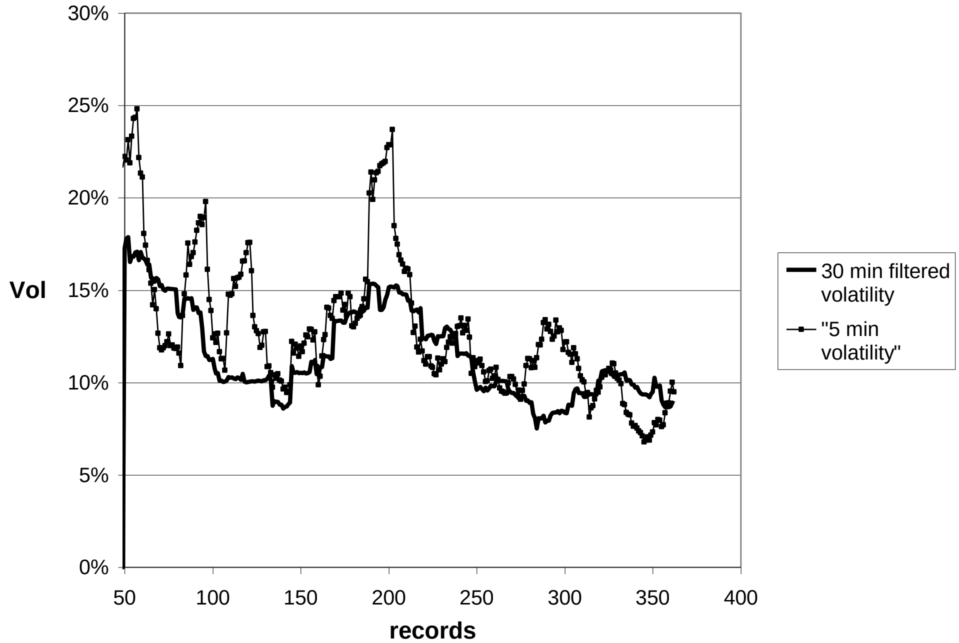


**Fig. 3.** Optimally filtered Nasdaq-100 volatility in the period between 04/26/2005 and 05/31/2005. Black curve with squares shows each sixth readout from the intraday volatility (compare to Fig. 2). The solid black curve shows the intraday volatility sampled according to Equations (8).

From Fig. 3 we observe that "optimally sampled" volatility generally follows the redacted volatility from the raw sample, i.e., the volatility computed from the sample with 30-minute frequency, but does not register sharp peaks, which we can attribute to microstructure noise. This exercise seems to indicate that the recipes provided by SMZ and Bandi-Russell are sound. Another set of the filtering of intraday results is given in Appendix B.

## 3. First empirical test of Ellsberg-ES asset pricing theory

To determine the elasticity of market volatility to the risk-free rate, we simply need to regress the logarithm of volatility on the risk-free rate, which I identify with an overnight Feds rate:

$$\log \Sigma_t = \alpha + \beta \cdot r_t + \varepsilon_t \qquad (9)$$

In Equation (9), $\Sigma_t$ is the Nasdaq volatility measured according to prescriptions in Section 2, $r_t$ is a return on a suitable proxy (Fed funds rate as *Proxy 1*; a synthetic rate formed from Fed funds rate and yield on T-bills with maturities between 1 and 14 weeks issued during April, 2005, as *Proxy 2*) and β is the required elasticity. The results of OLS regression (9) are provided in Table 2.

**Table 2. The sensitivity of filtered a Nasdaq volatility to the risk-free rate. Nasdaq-100 index volatility is determined as moving average volatility of 30-minute average returns between 04/26/2005 and 05/25/2005.** Fed funds rate (Proxy 1) is assumed to be constant throughout the day. Proxy 2, is a synthetic rate, formed from Federal funds rate and the primary yield on T-bills auctioned during that period. Standard deviations of regression parameters are given in parentheses. Both regression parameters are valid at P<0.1% for *Proxy 1* and the slope—for the *Proxy 2*. The intercept for the *Proxy 2* is valid at 5%. A very significant 25% of the variability of the index volatility can be ascribed to the change in a risk-free rate in the case of *Proxy 1*.

| Regression parameter | *Proxy 1* | *Proxy 2* |
|---|---|---|
| $\alpha$ | 5.236 (0.680) | 1.234 (0.482) |
| $\beta$ | -258.2 (23.93) | -117.3 (17.0) |
| Adj. $R^2$ | 24.9% | 11.9% |

We observe that the sensitivity of the index volatility to the risk-free rate has the correct, i.e., negative, sign and is, indeed, very high. How reasonable is this value? A simple differentiation of Equation (6) gives the following value for the elasticity:

$$\beta_{ES} = -\frac{2}{r} \tag{10}$$

Given that the Fed funds rate hovered around 2.9% in the described period, the estimate according to the Equation (10), gives $\beta_{ES} \approx -69$. This estimate is roughly 4 times lower than the estimation from the empirical data.

Equation (10) can be compared with an equally crude estimate from the DDM. Namely, if the index price understood as a combination of listed firms is expressed through a conventional formula:

$$P = \frac{Div}{k}$$

where $k$ is a required return on capital, the elasticity of the volatility to the risk-free rate will be expressed by an approximate formula:

$$\beta_{DDM} \approx -\frac{\sigma_{risk-free}}{k} \ll 1$$

Here, $\sigma_{risk-free}$ is the volatility of the risk free rate. This, preliminary set of empirical results seemingly supports the main qualitative feature of the Epstein-Schneider theory, namely, an extreme sensitivity of the index volatility to the movements of the risk-free rate. Yet, the final conclusion is disconcerting. The elasticity of the stock index to the risk-free rate below $\beta_{ES}$ can be explained away by the conventional mantra that not 100% of the return volatility is engendered by the Epstein-Schneider mechanism. According to this interpretation, Equation (10) can be considered as an upper limit for the elasticity in the ES. However, as we observe from Table 2, the daily-integrated empirical volatility is even more sensitive to the risk-free rate than the value inferred from Equation (10). Hence, no easy explanation suffices within the ES framework. Even if we surmise that the theory must account for the intrinsic leverage of an average company in the index—a currently unproven conjecture—a conventional rule of delevering (Copeland, Koller and Murrin, 1995) provides too high a value for the debt-to-asset ratio:

$$\hat{\beta}_{Levered,ES} = \beta \Big/ \left(1 - \frac{D}{V}\right)$$
$$D/V \approx 75\% \qquad (11)$$

The current test should not be construed as a refutation of the Epstein-Schneider theory. First, the regressed sample of daily integrated volatilities is too small (25 days). Second, we did not include other possible explanatory variables such as the expected market volatility or volume.

## 4. The test of elasticity of intraday volatility

To clarify the issue, I perform another set of tests. These tests involve the intraday regression:

$$\log(\Sigma_t) = \alpha_0 \; + \; \alpha_1 r_{t-1} + \alpha_2 \, r_{VXN,t-1day} + \alpha_3 \, \Delta V_{t-1day} + \epsilon_t \qquad (12)$$

Here $\Sigma_t = \widehat{\sigma}_t^2$ is a suitable proxy for instantaneous integrated volatility, $r_{t-1}$ is a Fed Funds rate in the previous integrated period of time, $r_{VXN,t-1day}$ is a past day's return on a volatility index (VXN for Nasdaq-100) and $\Delta V_{t-1day}$ is a past day's surprise in traded volume of the index per million of traded shares (activity in the said period was typically between 0.5 and 1 billion).

I choose two proxies for instantaneous volatility. Mechanics of filtering by reduction has been described in Section 2. For the period surrounding downgrade of the US AAA credit rating by Standard and Poors (August 5, 2011), the Equations (A7-A8) provide 12-minute interval for optimum sampling (rounded to the nearest minute). Fed Funds rate is reported round-the-clock but with 29-minutes intervals and I use a reading for the nearest past minute as a proxy for the instantaneous risk-free rate. Hence, the baseline level of a risk free rate is on average observed within the nearest hour of the volatility data collection. As a proxy for the instantaneous integrated volatility, I use two comparable choices. First, it is volatility from the GARCH(1,5) approximation with q=5 chosen as to correspond to the 1 hour period after past reading of the Fed Funds rate. Second, it is a moving average of annualized square deviation for the past half-day of trading. For the period of three weeks between July 26 and August 18, 2011, with the S&P downgrade approximately in the middle of this period, these two measures of integrated volatility have correlation $\rho_{12}$=76%.

The results of regression of Equation (12) are given in the Table 3. Elasticity of intraday volatilities is still high for the regression of Equation (12) but not as superficially high as the daily volatility without VXN and volume dummies. Yet, it becomes of order of unity if one adds a calendar dummy indicating a "regime change" on Aug. 5, 2011 when S&P downgraded US Government debt.

**Table 3. Factors determining intraday logarithmic integrated volatility in the period between July 26 and August 17, 2011**. The standard deviations are given in parentheses. The factors valid at P<0.1% are boldfaced. The factor valid at P<1% is marked with asterisk. The values of regression parameters are comparable for both proxies. The explanatory value of regression of Equation (12) is comparable to the Equation (9). Coefficients for both integrated volatility proxies are comparable in magnitude. A calendar dummy is a variable that indicates a regime change on August 5, 2011 due to the downgrade of the US Treasury debt.

| *Regression factor* | *GARCH(1,5)* | *Moving average* | *GARCH(1,5)* | *Moving average* |
|---|---|---|---|---|
| $\alpha_0$, Intercept | **-0.8976** | **-1.219** | **-1.2084** | **-1.6536** |
| | **(0.0186)** | **(0.0335)** | | **(0.0671)** |
| | | | **(0.0362)** | |
| $r_{t-1}$, Fed Funds | **-5.492** | **-7.972** | -1.720* | -2.692* |
| | **(0.4066)** | **(0.7320)** | (0.5423) | |
| | | | | (1.003) |
| $r_{\_VXN}$, Volatility index | 0.3288* | 0.3578 | 0.04710 | -0.03644 |
| | (0.1138) | (0.2049) | (0.1105) | (0.2043) |
| ΔV, Volume change | **0.0005596** | **0.0009889** | **0.0008825** | **0.001441** |
| | **(0.000131)** | **(0.000234)** | **(0.000126)** | **(0.000234)** |
| Calendar dummy | **-** | **-** | **0.3550** | **0.4968** |
| | | | **(0.0363)** | **(0.06715)** |
| $R^2$ | 24.0% | 16.8% | 32.9% | 22.9% |

This does not necessarily mean invalidation of ES—vice versa, it might signify that a sudden change in elasticity of volatility in response to the risk free rate may be a predictor of important macroeconomic events.

## Conclusion

An empirical analysis below seems to prove high elasticity of the Nasdaq-100 market index with respect to the changes in risk free rate. This happens despite the fact that the contribution of the risk free rate to the required rate of return of the high tech companies is rather low. The theory of Epstein and Schneider is one candidate for explanation of this counterintuitive fact. Elasticity of the observed volatility with respect to the risk-free rate has a correct sign and correct order of magnitude. Statistical significance of elasticity disappears when other explanatory variable (breakpoint calendar dummy) is included. Yet, as is the case with all behavioral anomaly if more and more analysts use this coefficient to predict structural breaks, this effect might disappear altogether.

According to ES, the market indexes experience such a drastic change when the Fed adjusts the rate because no-dividend stock provides a rather small bias to distinguish a true noise from uncertainty. But if this explanation is true then the valuations of the stock market are not risk neutral, or even risk avoidant in the conventional sense. In particular, in the arbitrage of a small immediate payoff or a postponed very large one, a player in a lottery might prefer an illiquid ticket with potentially large but very uncertain rewards to the ready cash. ES explanation is that because the promised gain is certain—however unlikely—the degree of risk avoidance is very low.

Intuitive explanation of a very sharp dependence of the non-dividend stock's volatility on the risk-free rate is that, in the absence of dividend signaling, a very small cognitive bias ($\gamma$-dependent terms in the Equations (5) and (6)) being discounted in perpetuity produces a drastic change in valuation. In the presence of cognitive bias, the only information observed by investors is the (relatively small) asset volatility, which is grossly amplified by discounting it in perpetuity. The reassessment of all future cash flows with little unambiguous information and even a small observational bias creates a sizeable correction.

Another well-known anomaly that could be explained by ES is the fact that structural models of corporate debt require unrealistically high asset volatilities to explain empirically observed credit spreads (Huang and Huang, 2002, Duffie and Singleton, 2003). The stylized fact of the ES that price volatility of a stock can exceed the asset volatility of an underlying company by hundreds of times (Fig. 1) may help to reconcile structural credit risk models to the observed facts.

## Appendix A. The optimal filtering frequency of high-frequency microstructure noise

We outline the theory of recovering volatility from the microstructure noise according to Aït-Sahalia, Mykland and Zhang (2005) and Bandi and Russell (2006). To estimate the (integrated) return volatility V:

$$V = \sum_{j=1}^{M} r_j^2 \qquad \text{(A.1)}$$

where $r_j$ are returns, and $M$ is a sample size. Bandi and Russell consider the estimator for "raw," noisy returns: $\tilde{r}_j = r_j + \varepsilon_j$:

$$\hat{V} = \sum_{j=1}^{M} \left(r_j + \varepsilon_j\right)^2 = \sum_{j=1}^{M} r_j^2 + 2\sum_{j=1}^{M} r_j \varepsilon_j + \sum_{j=1}^{M} \varepsilon_j^2 \qquad \text{(A.2)}$$

where $\varepsilon_j$ is microstructure noise.

Under certain generic conditions, they prove the following theorems.

**Theorem 1** (Theorem 2 in Bandi and Russell (2005)).

*1. The first four moments of the noisy returns* $\tilde{r}_j$ *can be estimated through the equality:*

$$\lim_{M \to \infty} \frac{1}{M} \sum_{j=1}^{M} \tilde{r}_j^{\,q} - E_M\left(\varepsilon^q\right) \xrightarrow{\ p\ } 0 \qquad 1 \le q \le 4 \qquad \text{(A.3)}$$

*2. Second-order autocorrelation can be estimated through:*

$$\lim_{M \to \infty} \frac{1}{M-m} \sum_{j=m+1}^{M} \tilde{r}_j^{\,2} \tilde{r}_{j-m}^{\,2} - E_M\left(\varepsilon^2 \varepsilon_{-m}^2\right) \xrightarrow{\ p\ } 0 \qquad \text{(A.4)}$$

where $m<M$ is fixed. In Equation (A.4), $E_M[\varepsilon^2 \varepsilon_{-m}^2] = Cov[\varepsilon_j^2, \varepsilon_{j-m}^2]$, which we consider stationary, i.e., independent on the index j.

**Theorem 2** (Theorem 3 in Bandi and Russell (2005))

Bias of the noisy volatility estimator is equal to:

$$E_M\left(\hat{V} - V\right)^2 = M^2 \cdot E_M\left[\varepsilon^4\right] + 2\sum_{j=1}^{M} (M-j) E_M\left(\varepsilon^2 \varepsilon_{-j}^2\right) + 4E_M[\varepsilon^2] V +$$

$$2 \frac{h}{M} \left(Q + o(1)\right) \qquad \text{(A.5)}$$

In Equation (A.5), $Q$ remains finite for $M \to \infty$ and the fixed h, and is given by the following expression:

$$Q = \frac{M}{h} \sum_{j=1}^{M} \left( \int_{(j-1)\delta}^{j\delta} \sigma_u^2 du \right)^2 \qquad \text{(A.6)}$$

where h is a total duration of the observations and $\delta = h/M$ is a time between observations.

**Corollary 1**. Optimum sampling frequency, or equivalently for fixed h, optimum M, is determined by the cubic equation:

$$aM^3 + bM^2 + c = 0 \qquad \text{(A.7)}$$

with the coefficients a, b and c given by the following equations.

$$a = E_M[\varepsilon^4], \qquad b = \lim_{M\to\infty}\left(\sum_{j=1}^{M}\left(1-{}^{j}\!/_{M}\right)E_M[\varepsilon^2\varepsilon_{-j}^2]\right) \qquad c = -hQ \tag{A.8}$$

**Proof**. To minimize residual mean square error, we differentiate by M, the number of observations, the expression of Equation (A.5) for the bias:

$$\begin{aligned} {}^{\partial RMSE}\!/_{\partial M} &= {}^{\partial E_M\left(\hat{V}-V\right)^2}\!/_{\partial M} = 2M\cdot E_M\left[\varepsilon^4\right] + \\ &2\sum_{j=1}^{M}(1-j/M)E_M\left(\varepsilon^2\varepsilon_{-j}^2\right) + o(1) - 2\,{}^{h}\!/_{M^2}\left(Q+o(1)\right) = 0 \end{aligned} \tag{A.9}$$

Collecting $M$-dependent and neglecting o(1) terms, we obtain Equation (A.7). Note that, following Theorem 2, we assumed that the second term in (A.9) does not depend on $M$ for $M\to\infty$. Replacement of expectations in Equations (A.8-A.9) with their sample estimators in the expression for approximate positive root of the Equation (A.7) provides an Equation (8) in the text.

**Appendix B. Practical results of the filtering of the microstructure noise**

A histogram of differences between all returns and optimally sampled returns provides the best available pictorial representation of the spectrum of the microstructure noise to date. To equalize number of readings in the filtered and unfiltered sample, I used jackknifed (Efron and Tibshiriani, 1994, Shao and Tu, 1996) representation of the lower-frequency returns on the Nasdaq-100 index. The results of simulation are given in the following table. A sample portrait of the microstructure noise spectrum is given in Fig. 4.

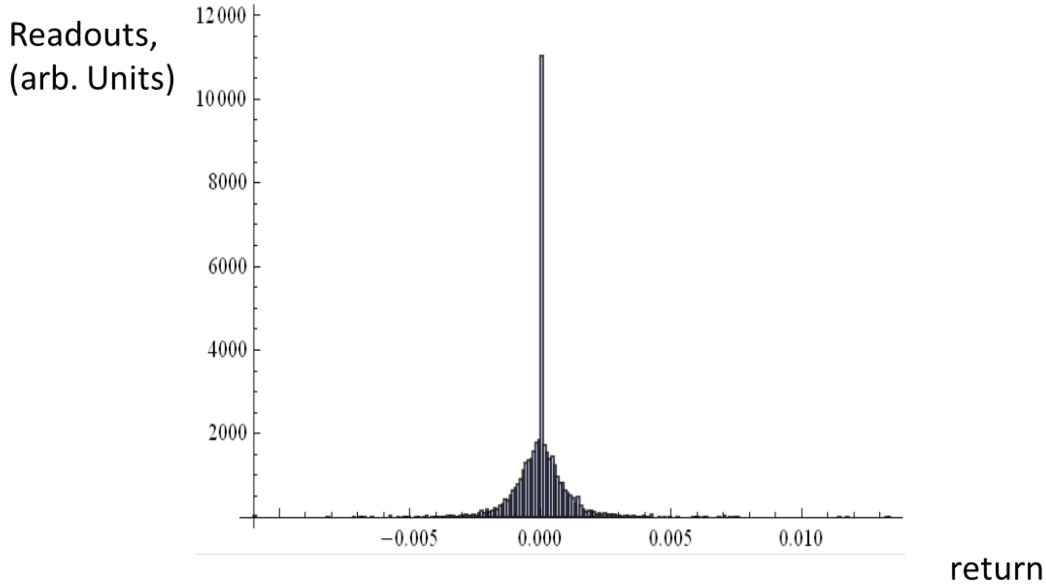


**Fig. 4.** The histogram of the returns of original jackknifed sample minus the sample filtered according to Bandi-Russell-Sahalia-Mykland-Zhang. It is natural to describe a high central peak as the result of microstructure noise such as rounded-off quotes and non-synchronous trading.

Table B1. Moments of the distribution of a proxy for “microstructure noise” (jackknifed differences between 5-minutes and 30-minutes return readings). As expected, the proxy for microstructure noise is unbiased and not skewed but has pronounced fat tails.

| *Scaled moment* | *Value* | *Note* |
|---|---|---|
| $m_1 /\sigma^{1/2}$ | -1.42E-03 | No bias |
| $m_2/\sigma$ | 1 | By construction |
| $m_3/\sigma^{3/2}$ | -0.0393 | No skew |
| $m_4/\sigma^2$ | 28.5 | Fat tails |

Another possible comparison includes of jackknifed redacted returns with the results of standard filtering using appropriate (for instance, Parzen or Epanechnikov, see Tsybakov, 2009) kernel.

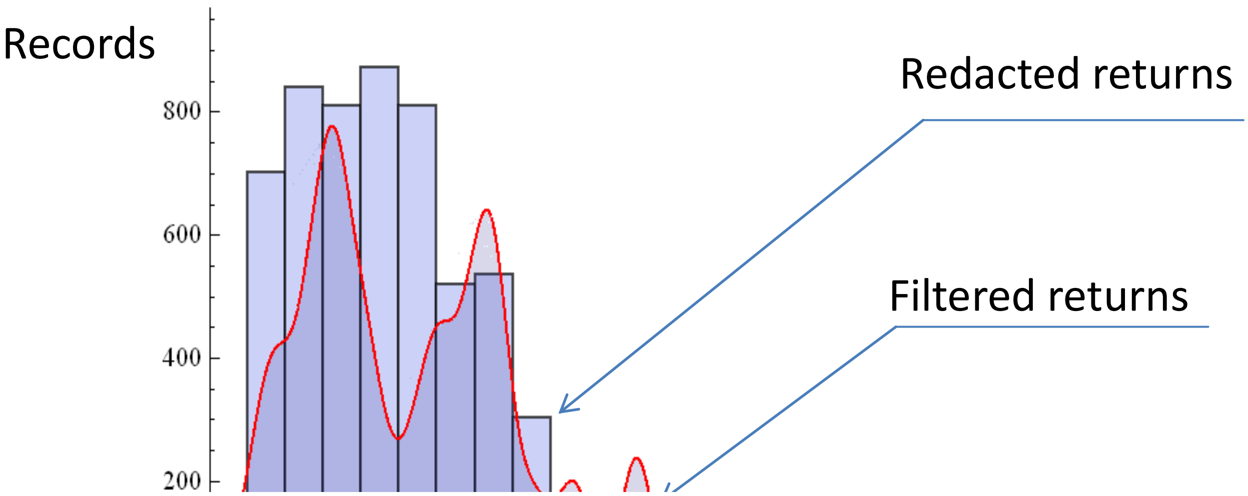


**Fig. 5** Histogram is the jackknifed distribution of Nasdaq-100 returns redacted according to the Bandi-Russell-Ait-Sahalia-Mykland-Zhang recipe (2005, 2006). The red curve is the histogram of actual returns smoothed down with a typical kernel.

As we observe from Fig. 5, filtering by omitting or averaging quotes on a time scale below optimum period is comparable to the standard filtering methods.

## REFERENCES

Andersen, Torben, Tim Bollerslev, Francis X. Diebold and Paul Labys (2003), Modeling and Forecasting Realized Volatility, *Econometrica*, 71(2): 579-625.

Bandi, F. M. and J. R. Russell (2005) Microstructure noise, realized variance and optimal sampling, Working Paper: Chicago, IL.

Bandi, F. M. and J. R. Russell (2006), Separating microstructure noise from volatility, *Journal of Financial Economics*, 79, 655-693.

Copeland, T., T. Koller and J. Murrin (1995), *Valuation: Measuring and Managing the Value of Companies*, Wiley: New York, NY.

Duffie, D. and K. J. Singleton (2003), *Credit Risk: Pricing, Measurement and Management*, Princeton University Press: Princeton, NJ.

Efron, B. and Tibshirani, R.J. (1994) *An Introduction to Bootstrap*. Boca Raton, FL: Chapman & Hall/CRC.

Fox, Craig R.; Tversky, Amos (1995), Ambiguity Aversion and Comparative Ignorance, *Quarterly Journal of Economics* 110 (3), 585–603.

Glosten, L. and P. Milgrom (1985), Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics*, 14, 71-100.

Hasbrouck, J. (2007) *Empirical Market Microstructure*. Oxford University Press: New York, NY.

Huang J., and M. Huang (2002), How much of the corporate-treasury yield spread is due to credit risk? A new calibration approach, Stanford University working paper: Stanford, CA.
ES (2008), Epstein, L. and M. Schneider, Ambiguity, Information Quality and Asset Pricing, *Journal of Finance*, 63, 197-228.

Kyle, A. S. (1985), Continuous auctions and insider trading, *Econometrica*, 53, 1315-1336.

Lerner, P., (2009), *Microstructure and Noise in Financial Markets*, VDM Verlag: Saarbrucken, Germany.

Mas-Colell A., M. Whinston and J. Green (1995), *Microeconomic Theory*, Oxford University Press: Oxford, UK.

McKay, C., *Extraordinary Popular Delusions and the Madness of Crowds,* Create Space: CA, 2011.
Percival, D. P. and A. T. Walden (2000), *Wavelet Methods for Time Series Analysis,* Cambridge University Press: Cambridge, UK.

Shao, J. and Tu, D. (1996) *The Jackknife and the Bootstrap*, New York, NY: Springer.

SMZ (2005), Aït-Sahalia, Y., P. *Mykland* and L. Zhang, How often to sample a continuous time process in the presence of market microstructure noise, *Review of Financial Studies*, 18, 351-415.

Tsybakov, A. (2009) *Introduction to Nonparametric Estimation*, Springer, Berlin: Germany.

ZMS (2005), Zhang, L., P. Mykland and Y. Aït-Sahalia, A tale of two time scales: determining integrated volatility with noisy high-frequency data, *Journal of American Statistical Association*, 100, 1394-1411.