Conditional heteroscedasticity model for discrete high-frequency price changes. With application to IBM trades data.

DMITRI KOULIKOV*

School of Economics and Management
Department of Economics
University of Århus
Århus C, 8000, Denmark

phone: +45 89421577
e-mail: dmitri_k@ut.ee

This revision:

February 7, 2002

Abstract

In this paper we present conditional heteroscedasticity models for time-series of discrete price changes in high-frequency financial data. They combine tractability of observation-driven GARCH models of Bollerslev (1986) with the simplicity of the ordered probit/logit structure of Hausman, Lo and MacKinlay (1992). In contrast to the ACM model of Russel and Engle (1998) and the ADS decomposition model of Rydberg and Shephard (1999), we separate groups of parameters driving conditional mean and conditional variance of the data, allowing us to test the effects of explanatory variables separately on the two moments of high-frequency price changes. We introduce two models belonging to the class outlined above: IV-GARCH model with short-memory volatility dynamics and IV-FIARCH model with long-range dependence in the conditional volatility. Application of the models to IBM trades dataset is provided.

JEL CLASSIFICATION: C22, C25, C51, G10
KEYWORDS: High-frequency financial data, Time-series of discrete random variables, Conditional heteroscedasticity, Markov chains, Non-linear econometric models.

*The author is grateful to the participants of the conference on "Market Microstructure and High-Frequency Data in Finance", Sandbjerg, Denmark, for their comments. The usual disclaimer applies.

Introduction

This paper presents a contribution to the literature on econometric modeling of high-frequency financial data. We introduce a class of observation-driven models for conditionally heteroscedastic discrete price changes in the spirit of GARCH models of Engle (1982) and Bollerslev (1986)[1]. This class includes both short- and long-memory models, where the latter is able to accommodate substantial persistence found in the volatility of discrete price changes. In addition to that, our models admit a relatively straightforward integration with financial duration models, such as the ACD model of Engle and Russel (1998) and Engle (2000), giving the framework for joint modeling of stochastically dependent inter-trade durations and price changes. Thus, this paper follows a research agenda put forward in Rydberg and Shephard (2000), where the authors propose compound Poisson process as the basic statistical model for high-frequency financial data.

Recent availability of high-frequency financial data, coupled with increased computing capacity, has spurred the literature seeking to develop a range of econometric techniques suitable for its statistical modeling. Among important recent contributions in the area are Davis, Rydberg, Shephard and Streett (2001), Engle (2000), Engle and Russel (1998), Gerhard and Pohlmeier (2000), Hausman, Lo and MacKinlay (1992), Rydberg and Shephard (1999, 2000), Russel and Engle (1998) and Engle (2000). A good survey of this relatively new econometric field is Hautsch and Pohlmeier (2001).

From a viewpoint of the established econometric methodology, high-frequency financial data presents several new challenges. First and foremost, high-frequency datasets contain collection of variables, such as bid and ask quotes, trade prices and trade volumes, were observations are separated by stochastic time intervals. As suggested by the market microstructure literature, these intervals themselves carry an important informational content and therefore should be modeled simultaneously with other variables. This point has been recently stressed in papers by Dufour and Engle (2000), Ghysels (2000) and Gerhard and Pohlmeier (2000).

Secondly, trade prices and quotes of many financial assets are discrete, reflecting the institutional structure of the markets. As documented in Rydberg and Shephard (2000), Campbell, Lo and MacKinlay (1997) and in section 3 of this paper, discrete price changes in high-frequency financial data exhibit statistical features similar to those normally observed in continuous low-frequency returns.

Finally, the real-time character of high-frequency data further contributes to its statistical complexity. Intra-day volatility patterns, news announcement effects and a range of market microstructure-specific idiosyncrasies make econometric modeling of this data particularly challenging.

In this paper we focus on discreteness of prices and quotes in high-frequency financial data. Our models are suitable both for price changes series "binned" into regular time intervals in the spirit of Davis, Rydberg, Shephard and Streett (2001), and for the original irregularly-spaced

[1] A comprehensive survey of the GARCH literature can be found in Bollerslev, Chou and Kroner (1992).

data. In the latter case, our models integrate with the class of stochastic duration models for financial data, providing a foundation for joint modeling of durations and price changes in high-frequency data.

Our work is motivated by the need for a tractable model that picks up essential dynamics of discrete price changes, in particular, of their second moment. Success of GARCH models of Engle (1982) and Bollerslev (1986) for low-frequency financial returns calls for the development of a similar observation-driven conditional heteroscedasticity model for discrete data. While both the Autoregressive Conditional Multinomial (ACM) model of Russel and Engle (1998) and the ADS decomposition model of Rydberg and Shephard (1999a) are able to describe the observed volatility clustering phenomenon in high-frequency price changes, neither of the two has a single underlying parameter driving the volatility. Among other things, this fact complicates testing of economically relevant hypothesis linked to the second moment of price changes in high-frequency data.

We reuse the idea of Hausman, Lo and MacKinlay (1992), whereby discrete distribution of price changes is approximated using the logistic distribution function in a framework resembling the ordered logit model. However, unlike Hausman, Lo and MacKinlay (1992), our models are cast entirely in terms of discrete random variables and therefore allow for simple estimation and diagnostic procedures. Moreover, probabilistic structure of the models lends itself to the study of stationarity and moments of the discrete price changes series.

Our models have two separate parameters: one driving conditional first moment of the discrete distribution of price changes, and the other driving its conditional variance. The latter is specified similarly to the GARCH model of Bollerslev (1986), allowing for a straightforward extension for the long-memory case and augmentation with a set of exogenous explanatory variables, such as announcement dummies and deterministic intra-day seasonality. In the paper we look in detail at both short- and long-memory cases, referred to as IV-GARCH and IV-FIARCH respectively[2].

The paper is organized as follows. Section 1 gives a short overview of the existing models for high-frequency financial data and discuss advantages and drawbacks of existing approaches. Section 2 introduces IV-GARCH and IV-FIARCH models, together with conditions for their existence and stationarity and an overview of the estimation and diagnostic methods. Section 3 gives the description of high-frequency IBM trades dataset used in the empirical part of the paper. Section 4 presents estimation results of the IV-GARCH and IV-FIARCH models. Conclusion summarizes the findings and discusses directions for the future research.

[2]These abbreviations stand for Integer-Valued Generalized Autoregressive Conditional Heteroscedasticity and Integer-Valued Fractionally Integrated Generalized Autoregressive Conditional Heteroscedasticity respectively.

1 Short overview of existing models for high-frequency financial data

In this section we give an overview of three existing models for discrete price changes in high-frequency financial data — ordered probit model of Hausman, Lo and MacKinlay (1992), ACM model of Russel and Engle (1998), and the ADS decomposition model of Rydberg and Shephard (1999) — and provide a brief discussion of a number of other contributions in the field. Some short comparison of the three models of price changes with our model is also given.

1.1 Ordered probit model of Hausman, Lo and MacKinlay (1992)

One of the first studies in the literature on modeling discrete price changes in high-frequency financial data, Hausman, Lo and MacKinlay (1992) propose an ordered probit model to capture essential features of such data. The ordered probit and logit models are well known from the cross-sectional econometrics, where its main application area includes modeling individual choices that have a natural ordering structure.

Hausman, Lo and MacKinlay (1992) motivate their model using the traditional framework for ordered probit and logit:

$$\Delta p_i^* = \mathbf{x}_i' \boldsymbol{\beta} + \epsilon_i \,,$$

where $\mathbf{x}_i$ is a vector of exogenous explanatory variables and ϵ_i is assumed to be an independent normal random variable with variance σ_i^2. In this setting Δp_i^* is an unobserved continuous state variable[3]. The observation rule is then given by:

$$\Delta p_i = \begin{cases} -K & \text{if} \quad \Delta p_t^* \in A_{-K} \\ -K+1 & \text{if} \quad \Delta p_t^* \in A_{-K+1} \\ \vdots & \vdots \quad \vdots \\ K & \text{if} \quad \Delta p_t^* \in A_K \end{cases} ,$$

where K denotes maximum allowable absolute price change, and $\{A_j\}_{j=-K}^{K}$ defines a finite partition of $\mathcal{R}$.

Hausman, Lo and MacKinlay (1992) make the variance of ϵ_i term dependent on a set of exogenous variables $\mathbf{w}_i$:

$$\sigma_i^2 = 1 + \mathbf{w}_i' \boldsymbol{\gamma} \,.$$

The need for heteroscedastic ϵ_i in their model is motivated by appealing to the diffusion models for prices of financial assets, where variance of increments depends on the sampling interval. Since price changes in high-frequency financial data are spaced irregularly, they include inter-trade durations in $\mathbf{w}_i$ to control for possible heteroscedasticity.

The models for Δp_i developed in section 2 have the structure superficially similar to that of ordered probit model of Hausman, Lo and MacKinlay (1992). In particular, our models inherit

[3] In the reminder of this paper we index observations in high-frequency financial data by subscript i. This convention underlines the fact that in general such data is not spaced in time regularly.

the ordered logit mechanism of modeling probabilities of price changes. However, the models developed in this papers do not have underlying continuous state variable Δp_i^* of Hausman, Lo and MacKinlay (1992)[4]. This difference is not purely motivational: variable Δp_i^* in the ordered probit model of Hausman, Lo and MacKinlay (1992) figures prominently throughout their paper, and in particular the diagnostic procedures developed by the authors rely on conditional independence and normality of Δp_i^*. This makes the specifications tests unnecessary complicated due to the latent character of the state variable.

Moreover, interpretation of Δp_i^* in the setting of financial markets is at best vague. Hausman, Lo and MacKinlay (1992) carefully avoid linking the state variable Δp_i^* to the hypothetical underlying continuous price, which is then rounded to the nearest tick by the ordering structure of their model. Their empirical findings indicate that the boundaries of $\{A_j\}_{j=-K}^{K}$ are misaligned with respect to the $\$\frac{1}{8}$ grid and vary substantially from stock to stock. However, when Δp_i^* does not have a price interpretation, the relevance of time varying σ_i^2 becomes unclear.

In this paper we show that the empirical success of the model in Hausman, Lo and MacKinlay (1992) lies in the flexibility of the ordered probit/logit structure in fitting the shape of the *discrete* distribution of high-frequency price changes. By looking at the ordered logit mechanism of modeling *discrete* Δp_i as merely an extremely convenient and parsimonious specification for the probability mass function of discrete price changes, we are able to cast an entire class of models in section 2 in terms of *discrete* random variables. We also briefly discuss other possible specifications for modeling time-series of discrete random variables suitable for high-frequency finance.

1.2 ACM model of Russel and Engle (1998)

Autoregressive conditional multinomial (ACM) model of Russel and Engle (1998) provides a general framework for modeling dynamics of the time series of discrete random variables. Russel and Engle (1998) specify a dynamic model for the discrete probability distribution of high-frequency price changes Δp_i. For price changes in transaction data the state space of Δp_i is assumed to be a bounded interval of $\mathcal{Z}$ symmetric around zero[5]. Let probabilities of individual price changes be denoted by $\boldsymbol{\pi} = (\pi_{-K}, \ldots, \pi_K)'$, where the support of Δp_i has the

[4]Historically, ordered probit and logit models were always build upon linear regressions with continuous disturbances and appropriate classifying observation rules. In medical and biological applications the underlying state variable is often interpreted as time or drug dosage, whereas in economics it most often represents level of utility. For an overview of the models refer to McFadden (1984). In this paper we work only with *discrete* random variables whose probability mass function is parametrized similarly to the structure of the ordered logit model; see section 2.

[5]We use this assumption throughout the paper. However, both the ACM model of Russel and Engle (1998) and the models developed in section 2 are not limited to this case.

length $2K+1$. Russel and Engle (1998) suggest a dynamic model for $\boldsymbol{\pi}_i$ of the following form:

$$
\begin{aligned}
h(\boldsymbol{\pi}_i) \quad = \quad & \sum_{j=1}^{p} A_j(\mathbf{1}_{\Delta p_{i-j}} - \boldsymbol{\pi}_{i-j}) \\
& + \sum_{j=1}^{q} B_j \mathbf{1}_{\Delta p_{i-j}} \\
& + \sum_{j=1}^{r} C_j h(\boldsymbol{\pi}_{i-j}) \\
& + G\mathbf{Z}_i \,,
\end{aligned}
\tag{1}
$$

where $\mathbf{1}_{\Delta p_i}$ denotes a $(2K+1) \times 1$ vector of the form $(\mathbf{1}_{\Delta p_i=-K}, \ldots, \mathbf{1}_{\Delta p_i=K})'$, $\mathbf{Z}_i$ stands for a set of weakly exogenous explanatory variables, and the link function $h(\cdot)$ is chosen such that the probabilities $(\pi_{-K}, \ldots, \pi_K)'$ sum up to unity. In their subsequent discussion of the model and its empirical applications Russel and Engle (1998) utilize multinomial logit specification for $h(\cdot)$.

With this level of generality almost every other model for the time series of discrete random variables will be a special case of the ACM model. In its practical application to the IBM transaction data Russel and Engle (1998) impose certain restrictions on the parameters of the ACM model in equation (1) justified by the considerations of response symmetry. These restrictions allow them to substantially reduce the number of parameters that have to be estimated.

While capable of producing very good fits into the high-frequency price changes data, ACM model has certain drawbacks from the empirical researcher perspective. Multinomial logit specification of the link function $h(\cdot)$ leads to difficulties in interpreting parameters of the model. In particular, direction in which included variables affect the probabilities of the states of Δp_i will not in general coincide with the signs of respective coefficients, and the total effect of an explanatory variable will depend on a subset of parameters. In the case of complicated dynamic specification, such as given in equation (1), numerical simulations from the model give a feasible solution to this interpretation problem. More generally, parameters of the ACM model do not naturally fall into groups driving conditional moments of the discrete distribution of Δp_i. While it is possible to identify a subset of parameters in the model that will enter the expressions for the conditional moments of Δp_i, is likely to be a complicated expression that is difficult to keep track of in practice. Therefore, testing of hypothesis that explicitly involve restrictions on the conditional moments is difficult.

The models for conditionally heteroscedastic discrete price changes presented in section 2 share many similarities with the ACM model of Russel and Engle (1998). However, several important differences are apparent. Firstly, our models are designed to have identifiable groups of parameters driving conditional first and conditional second moments of Δp_i. Hence, a range of economically interesting hypotheses related to the moments of discrete price changes can be tested directly, without the need to resort to post-estimation simulations. Secondly, by switching from the multinomial logit link function of the ACM model to the ordered logit or ordered probit link function, such as in the model of Hausman, Lo and MacKinlay (1992), substantial gains in terms of model parsimony are realized. This allows to reduce computational time needed to maximize the likelihood function — an important consideration in increasingly large high-frequency datasets. Finally, having a simpler specification for discrete price changes allows us to derive results pertaining to the stationary distribution of Δp_i.

1.3 ADS decomposition model of Rydberg and Shephard (1999)

Rydberg and Shephard (1999) propose a decomposition model for the discrete price changes in financial data, whereby Δp_i is assumed to be the product of three random processes as follows:

$$\Delta p_i = A_i \cdot D_i \cdot S_i \,,$$

where $\{A_i\}$ is the binary process on $\{0,1\}$ describing trading activity, $\{D_i\}$ is the binary process on $\{-1,1\}$ modeling direction of the price movement, and $\{S_i\}$ is the random process defined on the set of positive integers giving the size of the price change. All three processes are allowed to be interdependent with possible inclusion of other exogenous explanatory variables and intra-day seasonality. In their empirical analysis Rydberg and Shephard (1999) use the autologistic model for $\{A_i\}$ and $\{D_i\}$, and specify $\{S_i\}$ by the negative binomial GLARMA process.

ADS decomposition model has certain advantages from the point of view of market microstructure research and hypothesis testing. In many cases the market activity process $\{A_i\}$ can be the sole focus of research. The product of A_i and D_i represents the censored model of price movement, where Δp_i is at most allowed to change by one tick. Rydberg and Shephard (1999) report a number of interesting results concerning the dynamics and effect of exogenous variables on $\{A_i\}$, $\{D_i\}$ and $\{S_i\}$.

However, as in the ACM model of Russel and Engle (1998), the effects of explanatory variables in the ADS decomposition model are not directly tied to the moments of the discrete price changes. Although S_i is tightly related to the volatility of Δp_i, activity indicator A_i must also be accounted for in the implied second moment of the price changes distribution. The model for high-frequency price changes presented in this paper makes the link between parameter and the moments of Δp_i even more explicit.

1.4 Other contributions

Among other important contributions to the econometric modeling of high-frequency data we mention ACD-GARCH models for irregularly spaced financial data by Ghysels and Jasiak (1997), UHF-GARCH model by Engle (2000) and dynamic model for discrete bid-ask quotes with ARCH volatility by Hasbrouck (1999). The first two models are not designed to account for data discreteness, whereas Hasbrouck (1999) models it in the framework of the so-called rounding models of discreteness surveyed in Campbell, Lo and MacKinlay (1997) pp. 114-122.

ACD-GARCH model of Ghysels and Jasiak (1997) is based on the GARCH aggregation results of Drost and Nijman (1993), where aggregation intervals are stochastic and driven by the autoregressive conditional duration (ACD) model of Engle and Russel (1998). Ghysels and Jasiak (1997) introduce latent GARCH model that generates unobserved conditionally heteroscedastic returns at the highest observed frequency (normally 1 second). Parameters of this latent GARCH model are of the primary interest for the researcher. Observed irregularly spaced returns come from the aggregation of the latent data, where the aggregation intervals are stochastic and driven by the ACD model. This leads to the GARCH model with random coefficients that depend on the expected duration parameter from the ACD part. The basic model can be appended to include deterministic intra-day seasonality and effects of the exogenous explanatory variables. Ghysels and Jasiak (1997) present application of the ACD-GARCH to the IBM transaction dataset. They find that latent GARCH model features remarkably low volatility persistence — something that contrasts many other results, including those reported in Engle (2000). Ghysels and Jasiak (1997) interpret this results as the demonstration of the important role of the persistence of inter-trade durations, that together with the high-frequency returns create the evidence of substantial volatility persistence in high-frequency data.

ACD-GARCH model of Ghysels and Jasiak (1997) is one of the few contribution in the current literature attempting to link the volatility process of high-frequency returns with the process driving the inter-trade durations through their joint modeling. Another such attempt is made in Russel and Engle (1998), who also use the ACD model to fit the durations data. However, as also pointed out in Ghysels (2000), both model are reduced to the two-step framework, where the ACD model for durations data is estimated first under the assumption of exogeneity from the process driving the high-frequency returns or price changes data. Ghysels and Jasiak (1997) also attempt to test causality from the volatility of high-frequency returns to the inter-trade durations and find some supporting evidence for it.

UHF-GARCH model of Engle (2000) uses high-frequency returns scaled by the actual inter-trade durations for modeling in the usual GARCH framework. Scaling of the returns by the square root of durations is intuitively justified as the natural measure of the volatility per unit of time. Like Russel and Engle (1998), Engle (2000) also studies the effect of actual and predicted durations on the conditional second moment of scaled returns. However, in contrast to Russel and Engle (1998), Engle (2000) finds that actual durations have a statistically significant effect on the variance of scaled high-frequency returns.

Both the ACD-GARCH model of Ghysels and Jasiak (1997) and UHF-GARCH model of Engle (2000) ignore the inherent discreteness of the high-frequency financial data and fit traditional GARCH models into it. There have been no studies up to date discussing outcomes of this modeling decision on the performance of GARCH models. As documented in Hausman, Lo and MacKinlay (1992), Campbell, Lo and MacKinlay (1997) pp. 107-114, Russel and Engle (1998), Rydberg and Shephard (2000) and many other studies, high-frequency transaction data normally contains a large proportion of zero price changes and, therefore, zero returns. In addition to that, minimum price change of one tick is usually sufficiently coarse compared to the price level of the asset, leading to the bunching of high-frequency returns around the points of support of the discrete price change distribution; see Szpiro (1998) and Crack and Ledoit (1996) for more on this effect. For some distributional assumptions, such as GED in the

EGARCH model of Nelson (1991), the concentration of probability mass on the zero returns may lead to numerical instabilities and failures to estimate the model; see Hasbrouck (1999). Moreover, predictions from such models are likely to fail to generate sufficiently large amount of zero returns and to pick up the bunching.

Hasbrouck (1999) proposes a dynamic model for the discrete bid and ask quotes that is largely motivated by the insights from the market microstructure theory. Apart from the discreteness, his model features ARCH effects and incorporates costs of market making. Hasbrouck (1999) approach discreteness by the rounding of continuous data generated from the latent time-series process with conditionally heteroscedastic innovations. The rounding is asymmetric and is related to the cost of market making. This way of introducing discreteness into the model goes back to the contributions of Gottlieb and Kalay (1985) and Ball (1988), who use this setup to consider effects of discreteness on the estimator of variance of the continuous underlying process. Inference in the model is complicated by the presence of several latent components and is based on the recursive likelihood calculations, where the state variables are integrated out using numerical methods. Empirical findings of Hasbrouck (1999) imply highly peaked distribution of the innovations to the unobserved price process together with the relatively low degree of persistence of their variance.

2 IV-GARCH and IV-FIARCH models for high-frequency financial data

In this section we present a class of models for time-series of discrete conditionally heteroscedastic price changes in high-frequency financial datasets. The models are motivated by Hausman, Lo and MacKinlay (1992) ordered probit model for transactions data with the variance process similar to the GARCH model of Bollerslev (1986). The models belong to the class of observation driven models in the sense of Cox (1981) and lead to the straightforward maximum likelihood based inferential procedures. The basic specification can be also used to parametrize conditional variance of discrete price changes in terms of a background driving unobserved process.

2.1 Empirical observations on the distribution of high-frequency price changes in financial data

Before we proceed with discussion of IV-GARCH and IV-FIARCH models in the following subsections, we present several stylized facts pertaining to the statistical properties of high-frequency price changes. For a much broader survey of general empirical regularities of high-frequency financial data refer to Campbell, Lo and MacKinlay (1997) pp. 107-114. Here we emphasize three most notable features of Δp_i:

1. Figure 4 depicts marginal probability mass function of high-frequency price changes in IBM trades dataset. It is seen that the function is nearly symmetric around $\Delta p_i = 0$,

with high concentration of the probability mass on the middle state and almost no mass for $|\Delta p_i| \geq \$\frac{1}{2}$; see Campbell, Lo and MacKinlay (1997) pp. 107-114 and Hautsh and Pohlmeier (2001) for the similar evidence in other high-frequency datasets.

2. Another notable regularity of the data is a significant negative first-order autocorrelation of $\{\Delta p_i\}_{i\geq 1}$ series; see Figure 5. It follows that conditional on the sign of previous observation, distribution of Δp_i will be asymmetric; see Figure 1 for the illustration. The negative autocorrelation is consistent with the bid-ask bounce model of Roll (1984).

FIGURE 1 IS ABOUT HERE

3. Finally, as it is the case of many financial series with continuous support, $\{\Delta p_i\}_{i\geq 1}$ appears to exhibit dynamic heteroscedasticity. In the case of IBM data it is illustrated by the correlogram of $|\Delta p_i|$ on Figure 5. We also demonstrate probability mass function of Δp_i on Figure 2 at two different trading dates within our sample, where the difference in the tail mass distribution is apparent.

FIGURE 2 IS ABOUT HERE

In the next subsections we present a framework for econometric modeling of discrete price changes, where the three properties outlined above are accounted for in a relatively parsimonious and mathematically tractable way. The models can be easily extended in many other directions, making them suitable for testing a range of market microstructure related theories. In sections 3 and 4 we evaluate the fit of the models on the real-world IBM trades dataset.

2.2 Discrete distribution for high-frequency price changes

The discrete distribution for Δp_i in high-frequency financial series forms the basic building block of the models proposed further in this section. While there is a large class of statistical distributions with discrete support — a good overview of these can be found in Feller (1968) and Johnston and Kotz (1969) — most are restricted to the non-negative counts and therefore are not suitable for modeling Δp_i. The few parametric discrete distributions that are defined for the intervals of $\mathcal{Z}$ do not seem to have enough flexibility to accommodate a range of patterns of Δp_i that was documented in the previous subsection. Below we introduce a parametrization for the discrete distribution of high-frequency price changes that allows us to pick up changes in the first and second moments of the discrete data using as few parameters as possible.

We model the set of probability atoms $\boldsymbol{\pi} = (\pi_{-K}, \ldots, \pi_K)$ on a bounded interval of $\mathcal{Z}$ symmetric around zero as a function of two parameters linked to the first two moments of Δp_i[6]. In the reminder of the paper the two parameters are denoted μ and σ^2, where we allow for non-linear relationship between μ and $\mathbb{E}(\Delta p_i|\mu, \sigma^2)$ and σ^2 and $\mathbb{V}(\Delta p_i|\mu, \sigma^2)$. In addition to that, the following assumptions are used later in the paper:

[6]Here and henceforth we use the same notation as in subsection 1.

Assumption 1 *Functions* $\pi_{-K}(\mu,\sigma^2)\dots\pi_K(\mu,\sigma^2)$ *satisfy:*

1. *For all* $k=-K\dots K$ $0<\pi_k(\cdot,\cdot)<1$ *and* $\sum_{k=-K}^{K}\pi_k(\cdot,\cdot)=1$.

2. *Function* $\pi_{-K}(\mu,\sigma^2)\dots\pi_K(\mu,\sigma^2)$ *have the following limits:*

$$\begin{aligned}\lim_{\sigma^2\to 0}\pi_k(\mu,\sigma^2) &= \delta \qquad \text{for } k=-K\dots,-1,1,\dots,K\\ \lim_{\sigma^2\to 0}\pi_0(\mu,\sigma^2) &= 1-\delta\\ \lim_{\sigma^2\to\infty}\pi_k(\mu,\sigma^2) &< 1\,,\end{aligned}$$

where $0\le\delta<1$ *is small.*

3. *For all* $k=-K\dots K$*, functions* $\sigma^2\mapsto\pi_k(\mu,\sigma^2)$ *are Lipschitz, s.t.* $|\pi_k(\mu,x)-\pi_k(\mu,y)|\le C_k|x-y|$.

The first assumption above imposes restriction on probabilities $\boldsymbol{\pi}$, ensuring that the probability mass remains at all points in the support of Δp_i, regardless of the values of μ and σ^2. This restriction is common in econometric models with absolutely continuous distributions and time-varying volatility, such as GARCH and EGARCH models. It also plays an important role in ensuring dynamic stability of the time-series models based on this discrete distribution.

The second assumption gives some natural restrictions on the behavior of $\boldsymbol{\pi}$ as the function of σ^2. First two limits guarantee that the probability mass concentrates at the middle point of the support, possibly leaving only some limited probability mass in the tails, as σ^2 approaches zero. The third limit ensures stability of the probability distribution as $\sigma^2\to\infty$ by requiring the functions $\pi_{-K}(\mu,\sigma^2)\ \dots\ \pi_K(\mu,\sigma^2)$ to have a well-defined limit.

The third assumption imposes extra regularity requirements on the functions $\pi_{-K}(\mu,\sigma^2)$ $\dots$ $\pi_K(\mu,\sigma^2)$. In particular, it guarantees their smoothness with respect to σ^2 — a desired property in an econometric model.

General framework presented above resembles the one outlined by Russel and Engle (1998), but some important differences are present. Notably, as was mentioned in subsection 1.2, the model for discrete distribution of Δp_i in this paper is designed to have identifiable groups of parameters associated with the moments of high-frequency price changes. Because volatility in finance plays a prominent role and many hypothesis are specifically linked to the second moment of financial data, the structure of our model should provide a powerful and convenient tool for the empirical research. Another important distinction of our approach is to make $\pi_{-K},\dots,\pi_K$ the functions of parameters of interest, rather than to model evolution of $\boldsymbol{\pi}$ in the generalized VAR framework. As will become more clear below, we sacrifice some flexibility of the generalized VAR framework of Russel and Engle (1998) in order to obtain more parsimonious, easier to interpret structure of the discrete distribution of Δp_i. In fact, for modeling price changes in high-frequency data, one hardly needs completely flexible parametrization of $\boldsymbol{\pi}$. As we saw in the previous subsection, the spectrum of distributions of Δp_i have clear common features, such as pronounced concentration of probability mass on zero price change and thin tails. We make use of these facts to simplify our model and to gain in its interpretability.

Once the suitable mapping $(\mu, \sigma^2) \mapsto \boldsymbol{\pi}$ is established, a parametrization of μ and σ^2 can be selected. For example, in the spirit of observation driven models of Cox (1981), μ and σ^2 can be made dependent on the history of the process and a set of exogenous variables. Another suggestion is to parametrize them in terms of latent state variables. Moreover, it becomes possible to model μ and σ^2 together with other endogenous variables of interest in the high-frequency dataset. We will explore some of these possibilities later in the paper.

In the reminder of this subsection we describe a particular parametrization of functions $\pi_{-K}(\mu, \sigma^2), \ldots, \pi_K(\mu, \sigma^2)$ such that Assumption 1 is satisfied and parameters μ and σ^2 are linked to the moments of discrete distribution of Δp_i. Our choice is similar to the popular ordered logit model, where logistic distribution function is used as a link function between μ and σ^2 and the probabilities $\boldsymbol{\pi}$. This choice was originally motivated by the ordered probit model of Hausman, Lo and MacKinlay (1992), but as we mentioned previously, our model is not in the class of ordered logit models. We use logistic link function for the parametrization of $\boldsymbol{\pi}$ because of its convenience in describing the variety of forms of the probability mass distributions of Δp_i observed in the data; refer to subsection 1.1 for the discussion[7].

Logistic distribution function is a continuous bounded function with two parameters: location parameter μ and scale parameter σ^2. In addition to these, probabilities $\pi_{-K}, \ldots, \pi_K$ are defined using an additional set of parameters $\boldsymbol{\alpha} = (-\alpha_{K-1}, \ldots - \alpha_1, -1, 1, \alpha_1, \ldots \alpha_{K-1})$ in the following way:

$$\begin{aligned} \pi_{-K}(\mu, \sigma^2) &= \frac{1}{1 + e^{-\frac{-\alpha_{K-1} - \mu}{\sigma}}} \\ &\vdots \\ \pi_0(\mu, \sigma^2) &= \frac{1}{1 + e^{-\frac{-1-\mu}{\sigma}}} - \frac{1}{1 + e^{-\frac{1-\mu}{\sigma}}} \\ &\vdots \\ \pi_K(\mu, \sigma^2) &= 1 - \frac{1}{1 + e^{-\frac{\alpha_{K-1} - \mu}{\sigma}}} . \end{aligned} \tag{2}$$

Together with the assumed structure of $\boldsymbol{\alpha}$, distribution of Δp_i is seen to be an interval of $\mathcal{Z}$ symmetric around 0. This parametrization ensures that the probabilities $\pi_{-K} \ldots \pi_K$ sum up to unity. Note that in the rest of this paper it will be assumed that parameters $(\alpha_1, \ldots \alpha_{K-1})$ are constants, probably unknown, not depending on the history of the process or any set of exogenous variables. Parameters $\boldsymbol{\alpha}$ can be thought of as defining marginal distribution of price changes, whereas the variety of conditional distributions of Δp_i seen in subsection 2.1 is modeled using parameters μ and σ^2.

Symmetric discrete distribution of Δp_i around its middle state, normally zero price change,

[7]Other models for heteroscedastic sequences of discrete random variables can be constructed. For example, consider partition of the unit interval $[0, 1)$ into subintervals $A_1 = (0, \alpha_1]$, $A_2 = (\alpha_1, \alpha_2]$ and $A_3 = (\alpha_2, 1]$ s.t. $0 < \alpha_1 < \alpha_2 < 1$. Assign $\pi_i = \lambda_1(A_i)$ for $i = 1, 2, 3$ where λ_1 is the Lebesgue measure on $[0, 1)$. By parametrizing α_1 and α_2, a variety of shapes of the trivariate distribution of Δp_i can be achieved. In particular, $\mathbb{E}(\Delta p_i)$ changes when A_2 is moved around the unit interval and $\lambda_1(A_2)$ stays constant, while $\mathbb{V}(\Delta p_i)$ varies depending on $\lambda_1(A_2)$.

obtains whenever location parameter μ is zero. When $\mu \neq 0$, probability mass swings to the left or to the right tail of the distribution, leading to non-zero expected price change. This gives convenient way of picking up variations in the conditional first moment of Δp_i in the real-world data. This mechanism is seen from the expression for the first moment of Δp_i given by:

$$\mathbb{E}(\Delta p_i|\mu, \sigma^2) = \sum_{k=-K}^{K} k \cdot \pi_k(\mu, \sigma^2),$$

where $\pi_{-K} \dots \pi_K$ are defined in equation (2) and unit of measurement of $\mathbb{E}(\Delta p_i)$ is the number of ticks by which the price is expected to change. As apparent from this equation, whenever parameters $\boldsymbol{\alpha}$ have the structure given above, probability atoms in the left and in the right tail of the distribution will be equal to each other whenever $\mu = 0$ and different otherwise.

The second moment of Δp_i is given by the following expression:

$$\mathbb{V}(\Delta p_i|\mu, \sigma^2) = \sum_{k=-K}^{K} k^2 \cdot \pi_k(\mu, \sigma^2) - \Big(\mathbb{E}(\Delta p_i|\mu, \sigma^2)\Big)^2.$$

The measurement unit of the variance is given by squared ticks. When $\mu = 0$, the second part of the expression drops out, and as follows from equation (2), the remaining tail probabilities in the expression are scaled proportionally to the parameter σ^2, linking it to $\mathbb{V}(\Delta p_i)$.

However, just as in most parametric discrete distributions, there will be cross effects of μ and σ^2 on $\mathbb{V}(\Delta p_i)$ and $\mathbb{E}(\Delta p_i)$ respectively, since $\pi_{-K}(\mu, \sigma^2) \dots \pi_K(\mu, \sigma^2)$ are functions of both parameters. Figure 3 plots first two moments of the simple trivariate model of Δp_i as the functions of μ and σ^2. For $\mathbb{E}(\Delta p_i)$, the increase in σ^2 leads to less pronounced dependence of the first moment on μ, and similar effect is observed for the dependence of $\mathbb{V}(\Delta p_i)$ on σ^2 when μ increases. The same result holds for models with larger support of Δp_i.

FIGURE 3 IS ABOUT HERE

Given constant parameters μ, σ^2 and $\boldsymbol{\alpha}$ the static model presented in this subsection produces a sequence of i.i.d. homoscedastic discrete random variables $\{\Delta p_i\}_{i \geq 1}$. By specifying μ and σ^2 in terms of the set of exogenous parameters the variables in $\{\Delta p_i\}_{i \geq 1}$ will be independent but not necessarily identically distributed. In the following subsections we will show how to introduce dependence into the time series of high-frequency price changes such that the stylized facts presented in 2.1 can be modeled in an adequate way.

2.3 IV-GARCH model for heteroscedastic discrete Δp_i

Building upon the model for homoscedastic price changes introduced in the previous subsection, we now show how to model sequences of heteroscedastic Δp_i in high-frequency data. As documented in subsection 2.1, and as will be seen in section 3, there is a dynamic dependence in the real world Δp_i series in both the first and the second moments. We pick this up by borrowing ideas from GARCH models of Bollerslev (1986). In particular, we parametrize σ^2 in terms of its own lags and a set of exogenous variables, with disturbances given by a sequence of martingale difference (henceforth MD) innovations. μ is assumed to be non-dynamic, possibly depending on a set of exogenous variables.

The first model for dependent heteroscedastic sequence of price changes features short-memory dynamic structure in the second moment and is defined as follows:

$$\begin{aligned} \Delta p_i &\sim \pi_{-K}(\mu_i, \sigma_i^2; \boldsymbol{\alpha}), \dots \pi_K(\mu_i, \sigma_i^2; \boldsymbol{\alpha}) \\ \mu_i &= \boldsymbol{x}_i' \boldsymbol{\beta} \\ \sigma_i^2 &= \gamma_{0i} + \gamma_1 \sigma_{i-1}^2 + \gamma_2 \Big(\Delta p_{i-1}^2 - \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2) \Big) , \end{aligned} \tag{3}$$

where $\pi_{-K}(\mu, \sigma^2; \boldsymbol{\alpha}), \dots \pi_K(\mu_i, \sigma_i^2; \boldsymbol{\alpha})$ denotes discrete distribution introduced in the previous subsection with parameters μ_i, σ_i and $\boldsymbol{\alpha}$. $\mathbf{x}_i$ collects the set of exogenous variables normalized to have zero mean. γ_{0i} can optionally be parametrized as $\mathbf{z}_i'\boldsymbol{\delta}$ to include effects of intra-day seasonality, news announcements and additional exogenous variables. As apparent from (3), volatility process σ_i^2 has one autoregressive component and one driving MD innovation. In the reminder of this paper this model is referred to as IV-GARCH(1,1) .

$\{\sigma_i^2\}_{i \geq 1}$ process in the IV-GARCH(1,1) model above is driven by the innovation terms $\Delta p_i^2 - \mathbb{E}(\Delta p_i^2 | \sigma_i^2)$[8]. They have a natural interpretation of volatility "surprises", i.e. unexpected increases or decreases in the volatility of the discrete random variable Δp_i. The innovations $\Delta p_i^2 - \mathbb{E}(\Delta p_i^2 | \sigma_{i-1}^2)$ are by construction martingale differences, since $\mathbb{E}\Big(\Delta p_i^2 - \mathbb{E}(\Delta p_i^2 | \sigma_i^2) \Big| \mathfrak{F}_{i-1} \Big) = 0$, where $\mathfrak{F}_i$ stands for the information generated by the process up to time i. The sequence of squared price innovation $\{\Delta p_i^2\}_{i \geq 1}$ in the volatility process is clearly not i.i.d. and therefore requires careful treatment in proofs of the distributional results and statistical properties of $\{\sigma_i^2\}_{i \geq 1}$.

Below we give simple necessary and sufficient condition for positivity of the volatility process in (3).

Lemma 1 *Volatility process in IV-GARCH(1,1) model is positive iff* $\sigma^2 \mapsto \gamma_0 + \gamma_1 \sigma^2 - \gamma_2 \mathbb{E}(\Delta p^2 | \sigma^2) > 0$ *for all* $\sigma^2 > 0$.

PROOF Write the noise-free skeleton of the volatility process in (3) as:

$$\sigma_i^2 = \sum_{j=0}^{i-1} \gamma_1^j \Big(\gamma_0 - \gamma_2 \mathbb{E}(\Delta p_j^2 | \sigma_j^2) \Big) + \gamma_1^i \sigma_0^2 . \tag{4}$$

Now, recursive substitution into the equation above and positivity of $\sigma^2 \mapsto \gamma_0 + \gamma_1 \sigma^2 - \gamma_2 \mathbb{E}(\Delta p^2 | \sigma^2)$ shows that:

$$\sigma_1^2 = \gamma_0 - \gamma_2 \mathbb{E}(\Delta p_0^2 | \sigma_0^2) + \gamma_1 \sigma_0^2 > 0$$

[8] At this point the difference between our view of the model in equation (3) and the traditional ordered probit/logit literature is most visible. Since we do not have any underlying latent variable with absolutely continuous distribution in (3), the only random innovation in our model is the discrete Δp_i. Traditional approach may call for the inclusion of the squared unobserved continuous Δp_i^*, which leads to complications due to its latent character. Also, diagnostic procedures in traditional ordered probit/logit models require calculation of the expected underlying innovation; see Hausman, Lo and MacKinlay (1992). We base our specification tests directly on Δp_i.

$$
\begin{aligned}
\sigma_2^2 &= \Big(\gamma_0 - \gamma_2 \mathbb{E}(\Delta p_1^2|\sigma_1^2)\Big) + \gamma_1 \Big(\gamma_0 - \gamma_2 \mathbb{E}(\Delta p_0^2|\sigma_0^2)\Big) + \gamma_1^2 \sigma_0^2 \\
&= \Big(\gamma_0 - \gamma_2 \mathbb{E}(\Delta p_1^2|\sigma_1^2)\Big) + \gamma_1 \sigma_1^2 > 0 \\
&\vdots
\end{aligned}
$$

□

Choice of parameters γ_0, γ_1 and γ_2 to satisfy conditions in Lemma 1 is always possible because according to Assumption 1 function $\sigma^2 \mapsto \mathbb{E}(\Delta p^2|\sigma^2)$ is continuous and bounded between 0 and K^2. In addition to that, a simpler sufficient condition for positivity of the volatility process in IV-GARCH(1,1) model is given by $\sigma^2 \mapsto \gamma_0 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma^2) > 0$, which is seen from (4).

Before establishing a form of stochastic stability of IV-GARCH(1,1) model in subsection 2.4, we show deterministic stability of the noise-free skeleton of the model as given in equation (4). We also obtain the lower bound for the volatility process in the IV-GARCH(1,1) model. We need the following extra regularity condition:

Assumption 2 *Function $\sigma^2 \mapsto \gamma_0 + \gamma_1\sigma^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma^2)$ defines a contraction.*

Proposition 1 *Under Assumption 2, deterministic part of the volatility process in IV-GARCH(1,1) model defined by the recursion $\sigma_i^2 = \gamma_0 + \gamma_1\sigma_{i-1}^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma_{i-1}^2)$ is globally stable and has a unique limit point given by the solution of $\underline{\sigma}^2 = \gamma_0 + \gamma_1\underline{\sigma}^2 - \gamma_2 \mathbb{E}(\Delta p^2|\underline{\sigma}^2)$.*

PROOF We utilize the invariance principle given in Theorem 2.9 of Tong (1990). We have to show that mapping $\sigma^2 \mapsto \gamma_0 + \gamma_1\sigma^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma^2)$ is continuous and bounded, and that Lyapunov function exists for the recursion $\sigma_i^2 = \gamma_0 + \gamma_1\sigma_{i-1}^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma_{i-1}^2)$.

1. Continuity of the mapping $\sigma^2 \mapsto \gamma_0 + \gamma_1\sigma^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma^2)$ follows from Assumption 1, whereby function $\sigma^2 \mapsto \mathbb{E}(\Delta p^2|\sigma^2)$ is continuous. Boundedness is an immediate consequence of the parameter restrictions $\gamma_0, \gamma_2 > 0$, $0 < \gamma_1 < 1$ and non-negativity of the function $\sigma^2 \mapsto \mathbb{E}(\Delta p^2|\sigma^2)$.

2. Define an identity map $V(\sigma^2) \equiv \sigma^2$. In order for V to be Lyapunov function for the recursion $\sigma_i^2 = \gamma_0 + \gamma_1\sigma_{i-1}^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma_{i-1}^2)$, we have to check that:

   $$V(\gamma_0 + \gamma_1\sigma^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma^2)) - V(\sigma^2) = \gamma_0 + \gamma_1\sigma^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma^2) - \sigma^2 \leq 0$$

   for $\sigma^2 \in G \subseteq \mathcal{R}_+$. Under Assumption 2 there exists is a unique solution of the equation $\gamma_0 - (1-\gamma_1)\sigma^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma^2) = 0$ provided that $\lim_{\sigma^2 \downarrow 0} \mathbb{E}(\Delta p^2|\sigma^2) \leq \frac{\gamma_0}{\gamma_2}$. We denote this solution by $\underline{\sigma}^2$. Hence, V is Lyapunov function for the recursion $\sigma_i^2 = \gamma_0 + \gamma_1\sigma_{i-1}^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma_{i-1}^2)$ on $G = [\underline{\sigma}^2, \infty)$.

Global stability of the recursion $\sigma_i^2 = \gamma_0 + \gamma_1\sigma_{i-1}^2 - \gamma_2 \mathbb{E}(\Delta p^2|\sigma_{i-1}^2)$ follows from the fact that it converges to $\underline{\sigma}^2$ from any $\sigma_0^2 \in G$. □

It is also seen that the expected value of the volatility process in equation (3) is given by:

$$\mathbb{E}(\sigma_i^2) = \frac{\mathbb{E}(\gamma_{0i})}{1-\gamma_1}.$$

2.4 Existence and uniqueness of stationary distribution of IV-GARCH(1,1) model

In this subsection we prove existence and uniqueness of the stationary distribution of IV-GARCH(1,1) model using results from the literature on Markov chains models in general state spaces. In particular, we will follow the line of proofs given in Davis, Rydberg, Shephard and Streett (2001). IV-GARCH(1,1) model and the CBIN model introduced by these authors have similar probabilistic structures, and the cited work have been very helpful in establishing stationarity results for our model.

We begin the proof by presenting two auxiliary results from Glynn and Meyn (1997) that will be used later in this subsection:

Corollary 1 *Suppose that there exists a measurable function $V : \mathcal{X} \mapsto [0, \infty)$ and a set $A \in \mathfrak{B}(\mathcal{X})$ satisfying:*

1. *For some $b < \infty$,*
$$PV \leq V - 1 + b\mathbf{1}_A \,.$$
2. $$\lim_{i\to\infty} \sup_{a\in A} \mathbb{E}\Big(V(\sigma_i^2)\mathbf{1}(\tau_A > i) \Big| \sigma_0^2 = a \Big) = 0 \,.$$
3. *For each $m \geq 1$, the family of probability measures $\{\frac{1}{m}\sum_{k=1}^{m} P^k(a,\cdot) : a \in A\}$ is tight.*

Then the chain is bounded in probability on average.

Theorem 1 *If $\{\sigma_i^2\}_{i\geq 1}$ is a weak Feller, then for each $m \geq 1$ and compact A, the family of probability measures $\{\frac{1}{m}\sum_{k=1}^{m} P^k(a,\cdot) : a \in A\}$ is tight.*

Proof of Corollary 1 is given in Glynn and Meyn (1997) and proof of Theorem 1 can be found in Davis, Rydberg, Shephard and Streett (2001).

Note that the volatility process in IV-GARCH(1,1) model defines a Markov chain on the positive half-line:

$$\sigma_i^2 = \gamma_0 + \gamma_1 \sigma_{i-1}^2 - \gamma_2 \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2) + \gamma_2 \Delta p_{i-1}^2 \,, \tag{5}$$

where $\{\Delta p_i^2\}_{i\geq 1}$ is the sequence of non-i.i.d. innovations. We prove the following results:

Theorem 2 *Markov chain $\{\sigma_i^2\}_{i\geq 1}$ with transition function defined by (5) possesses unique stationary distribution under Assumptions 1 and 2.*

Proof To show existence and uniqueness of the stationary distribution of IV-GARCH(1,1) model we show that Markov chain $\{\sigma_i^2\}_{i\geq 1}$ defined by this equation is bounded in probability on average, is an e-chain and possesses a reachable state.

1. To show that the chain $\{\sigma_i^2\}_{i\geq 1}$ is bounded in probability on average we verify three conditions given in Corollary 1 of Glynn and Meyn (1997).

(a) Let function V be given by the identity map $V(x) = x$, let set A be an interval $[\underline{\sigma}^2, \frac{1+\gamma_0}{1-\gamma_1}]$, where $\underline{\sigma}^2 = \frac{\gamma_0 - \gamma_2 \mathbb{E}(\Delta p^2 | \underline{\sigma}^2)}{1-\gamma_1}$ is given in Proposition 1, and let $b = 1 + \gamma_2 \mathbb{E}(\Delta p^2 | \underline{\sigma}^2)$. Recall that (5) is defined on $\mathcal{R}_+$. We are required to show that $PV - V \leq -1 + b\mathbf{1}_A$. This follows from:

$$\begin{aligned} PV - V &\equiv \mathbb{E}(\sigma_i^2 | \sigma_{i-1}^2 = x) - x \\ &= \gamma_0 + \gamma_1 x + \gamma_2 \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2) - \gamma_2 \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2) - x \\ &= \gamma_0 + x(\gamma_1 - 1) \\ &\leq -1 + (1 + \gamma_2 \mathbb{E}(\Delta p^2 | \underline{\sigma}^2))\mathbf{1}_A \,. \end{aligned}$$

(b) Let set A be as before and note that be Cauchy-Schwartz inequality we have that:

$$\lim_{i\to\infty} \sup_{a\in A} \mathbb{E}\Big(\sigma_i^2 \mathbf{1}(\tau_A > i) \Big| \sigma_0^2 = a\Big) \leq \lim_{i\to\infty} \sup_{a\in A} \mathbb{E}^{\frac{1}{2}}\Big((\sigma_i^2)^2 \Big| \sigma_0^2 = a\Big) \mathbb{P}^{\frac{1}{2}}(\tau_A > i | \sigma_0^2 = a) \,. \tag{6}$$

The first term in the inequality above can be written as:

$$\begin{aligned} \mathbb{E}\Big((\sigma_i^2)^2 \Big| \sigma_0^2 = a\Big) &= \mathbb{E}\Big((\gamma_0 + \gamma_1 \sigma_{i-1}^2 + \gamma_2[\Delta p_{i-1}^2 - \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2)])^2 \Big| \sigma_0^2 = a\Big) \\ &= \mathbb{E}\Bigg(\mathbb{E}\Big(\gamma_0^2 + \gamma_1^2 (\sigma_{i-1}^2)^2 + \gamma_2^2[\Delta p_{i-1}^2 - \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2)]^2 \\ &\quad + 2\gamma_0\gamma_1\sigma_{i-1}^2 + 2\gamma_0\gamma_2[\Delta p_{i-1}^2 - \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2)] \\ &\quad + 2\gamma_1 \sigma_{i-1}^2 \gamma_2 [\Delta p_{i-1}^2 - \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2)] \Big| \sigma_{i-1}^2\Big) \Bigg| \sigma_0^2 = a\Bigg) \\ &= \gamma_0^2 + \gamma_1^2 \mathbb{E}\Big((\sigma_{i-1}^2)^2 \Big| \sigma_0^2 = a\Big) + \gamma_2^2 \mathbb{V}(\Delta p_{i-1}^2 | \sigma_0^2 = a) \\ &\quad + 2\gamma_0\gamma_1 \mathbb{E}(\sigma_{i-1}^2 | \sigma_0^2 = a) \,. \end{aligned}$$

By recursively substituting $\mathbb{E}\Big((\sigma_i^2)^2 | \sigma_0^2 = a\Big)$ into the expression above the following equation obtains:

$$\begin{aligned} \mathbb{E}\Big((\sigma_i^2)^2 \Big| \sigma_0^2 = a\Big) &= \gamma_0^2 \sum_{j=0}^{t-1} \gamma_1^{2j} + \gamma_1^{2t} a^2 + 2\gamma_0\gamma_1 \sum_{j=0}^{i-1} \gamma_1^{2j} \mathbb{E}(\sigma_{i-j-1}^2 | \sigma_0^2 = a) \\ &\quad + \gamma_2^2 \sum_{j=0}^{i-1} \gamma_1^{2j} \mathbb{V}(\Delta p_{i-j-1}^2 | \sigma_0^2 = a) \,. \end{aligned} \tag{7}$$

In this expression, $\mathbb{E}(\sigma_i^2 | \sigma_0^2 = a)$ can be written as follows:

$$\begin{aligned} \mathbb{E}(\sigma_i^2 | \sigma_0^2 = a) &= \mathbb{E}\Big(\gamma_0 + \gamma_1 \sigma_{i-1}^2 + \gamma_2[\Delta p_{i-1}^2 - \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2)] \Big| \sigma_0^2 = a\Big) \\ &= \mathbb{E}\Big(\mathbb{E}\Big(\gamma_0 + \gamma_1 \sigma_{i-1}^2 + \gamma_2[\Delta p_{i-1}^2 - \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2)] \Big| \sigma_{i-1}^2\Big) \Big| \sigma_0^2 = a\Big) \\ &= \gamma_0 + \gamma_1 \mathbb{E}(\sigma_{i-1}^2 | \sigma_0^2 = a) \\ &= \ldots = \gamma_0 \sum_{j=0}^{i-1} \gamma_1^j + \gamma_1^i a \,. \end{aligned}$$

From this we have that:

$$\begin{aligned} 2\gamma_0\gamma_1 \sum_{j=0}^{i-1} \gamma_1^{2j} \mathbb{E}(\sigma_{i-1}^2 | \sigma_0^2 = a) &= 2\gamma_0\gamma_1 \sum_{j=0}^{i-1} \gamma_1^{2j} \Big(\gamma_0 \sum_{l=0}^{i-j-2} \gamma_1^l + \gamma_1^{i-j-1} a \Big) \\ &= 2\gamma_0\gamma_1^i a \sum_{j=0}^{i-1} \gamma_1^j + 2\gamma_0^2\gamma_1 \sum_{j=0}^{i-1} \gamma_1^{2j} \sum_{l=0}^{i-j-2} \gamma_1^l \,, \end{aligned}$$

from where using the fact that $\sum_{l=0}^{i-j-2} \gamma_1^l \leq \frac{1}{1-\gamma_1}$ we arrive at the limit:

$$\lim_{i \to \infty} 2\gamma_0\gamma_1 \sum_{j=0}^{i-1} \gamma_1^{2j} \mathbb{E}(\sigma_{i-j-1}^2 | \sigma_0^2 = a) \leq \frac{2\gamma_0^2\gamma_1}{(1-\gamma_1)(1-\gamma_1^2)} \,.$$

Next, consider $\mathbb{V}(\Delta p_i^2 | \sigma_0^2 = a)$ in equation (7). Because the support of the random variable Δp_i^2 is a finite subset of $\mathcal{N}_0$, its variance will always be bounded. Let the upper bound of $\mathbb{V}(\Delta p_i^2 | \sigma_0^2 = a)$ be given by $\bar{V}$. Then we have the following limit:

$$\lim_{i \to \infty} \gamma_2^2 \sum_{j=0}^{i-1} \gamma_1^{2j} \mathbb{V}(\Delta p_{i-j-1}^2 | \sigma_0^2 = a) \leq \lim_{i \to \infty} \gamma_2^2 \sum_{j=0}^{i-1} \gamma_1^{2j} \bar{V} = \frac{\gamma_2^2 \bar{V}}{1-\gamma_1^2} \,.$$

Combining these results for the first part of equation (6) we get the following inequality:

$$\mathbb{E}\Big((\sigma_i^2)^2 \Big| \sigma_0^2 = a \Big) \leq \frac{\gamma_0^2}{1-\gamma_1^2} + a^2 + \frac{2\gamma_0^2\gamma_1}{(1-\gamma_1)(1-\gamma_1^2)} + \frac{\gamma_2^2\bar{V}}{1-\gamma_1^2} = c_1 < \infty \,.$$

The second part of equation (6) follows from the Theorem 11.3.4 of Meyn and Tweedie (1993), whereby:

$$\mathbb{P}(\tau_A > i | \sigma_0^2 = a) \leq \frac{\mathbb{E}(\tau_A | \sigma_0^2 = a)}{i+1} \leq \frac{V(a) + b\mathbf{1}_A(a)}{i+1} \leq \frac{a + 1 + \gamma_2 \mathbb{E}(\Delta p^2 | \underline{\sigma}^2)}{i+1} \,.$$

The second condition of the Corollary 1 then follows from:

$$\begin{aligned} \limsup_{i \to \infty} \sup_{a \in A} \mathbb{E}\Big(\sigma_i^2 \mathbf{1}(\tau_A > i) \Big| \sigma_0^2 = a \Big) &\leq c_1^{\frac{1}{2}} \lim_{i \to \infty} \sup_{a \in A} \left(\frac{a + 1 + \gamma_2 \mathbb{E}(\Delta p^2 | \underline{\sigma}^2)}{i+1} \right)^{\frac{1}{2}} \\ &\leq c_1^{\frac{1}{2}} \lim_{i \to \infty} \left(\frac{\frac{1+\gamma_0}{1-\gamma_1} + 1 + \gamma_2 \mathbb{E}(\Delta p^2 | \underline{\sigma}^2)}{i+1} \right)^{\frac{1}{2}} = 0 \,. \end{aligned}$$

(c) The third condition of the Corollary 1 is the consequence of Theorem 1 if we show that chain $\{\sigma_i^2\}_{i \geq 1}$ is a weak Feller chain. Recall that a Markov chain is said to be weak Feller if its transition function $P(\cdot, O)$ is a lower semicontinuous function for any open set $O \in \mathfrak{B}(\mathcal{X})$; refer to Tweedie (1998). Rewrite (5) as:

$$\sigma_i^2 = \gamma_0 + \gamma_1 \sigma_{i-1}^2 - \gamma_2 \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2) + \gamma_2 \Delta p_{i-1}^2 \,.$$

It follows that the Markov transition kernel $P(\cdot, O)$ for the chain defined by the equation above is given by:

$$P(x, O) = \sum_{k=-K}^{K} \mathbf{1}_O(\gamma_0 + \gamma_1 x - \gamma_2 \mathbb{E}(\Delta p^2 | x) + \gamma_2 k^2) \pi_k(x) \,.$$

Recall that function $\mathbf{1}_O$ is a lower semicontinuous function for an open set O, and that according to the Assumption 1 functions $x \mapsto \mathbb{E}(\Delta p^2|x)$ and $x \mapsto \pi_k(x)$, $k = -K \dots K$ are both continuous in x. Hence, $x \mapsto P(x, O)$ is a lower semicontinuous for the IV-GARCH(1,1) model.

This finishes the proof that the chain $\{\sigma_i^2\}_{i\geq 1}$ is bounded in probability on average.

2. Recall that a Markov chain is said to be an e-chain if the collection of Markov transition kernels $\{P^n f : n \geq 1\}$ is equicontinuous for each continuous function f with compact support; refer to Meyn and Tweedie (1993). Therefore, it is necessary to show that for any x, y in the state-space of the chain $\{\sigma_i^2\}_{i\geq 1}$ and for a given $\epsilon_1 > 0$ there is $\epsilon_2 > 0$ s.t. $|P_x^n f - P_y^n f| < \epsilon_1$ whenever $|x - y| < \epsilon_2$ for all $n \geq 1$.

   We start with one-step transition probabilities. Since by assumption f will be uniformly continuous and bounded, assume without loss of generality that $|f| \leq 1$. Observe that:

$$
\begin{aligned}
|P_x f - P_y f| &= \Bigg| \sum_{k=-K}^{K} f(\gamma_0 + \gamma_1 x - \gamma_2 \mathbb{E}(\Delta p^2|x) + \gamma_2 k^2)\pi_k(x) \\
&\quad - \sum_{k=-K}^{K} f(\gamma_0 + \gamma_1 y - \gamma_2 \mathbb{E}(\Delta p^2|y) + \gamma_2 k^2)\pi_k(y) \Bigg| \\
&= \Bigg| \sum_{k=-K}^{K} \Big(f(\gamma_0 + \gamma_1 x - \gamma_2 \mathbb{E}(\Delta p^2|x) + \gamma_2 k^2) \\
&\quad - f(\gamma_0 + \gamma_1 y - \gamma_2 \mathbb{E}(\Delta p^2|y) + \gamma_2 k^2) \Big) \pi_k(x) \\
&\quad + \sum_{k=-K}^{K} f(\gamma_0 + \gamma_1 y - \gamma_2 \mathbb{E}(\Delta p^2|y) + \gamma_2 k^2) \Big(\pi_k(x) - \pi_k(y) \Big) \Bigg| \\
&\leq \sum_{k=-K}^{K} \Big| f(\gamma_0 + \gamma_1 x - \gamma_2 \mathbb{E}(\Delta p^2|x) + \gamma_2 k^2) \\
&\quad - f(\gamma_0 + \gamma_1 y - \gamma_2 \mathbb{E}(\Delta p^2|y) + \gamma_2 k^2) \Big| \pi_k(x) \\
&\quad + \sum_{k=-K}^{K} \Big| \pi_k(x) - \pi_k(y) \Big| .
\end{aligned}
$$

   Now, using uniform continuity of f, the differences $|f(\gamma_0 + \gamma_1 x - \gamma_2 \mathbb{E}(\Delta p^2|x) + \gamma_2 k^2) - f(\gamma_0 + \gamma_1 y - \gamma_2 \mathbb{E}(\Delta p^2|y) + \gamma_2 k^2)|$ can be made less that $\epsilon' > 0$ whenever $|x - y| < \epsilon_2$ for any x, y in the state-space of the chain. Also recall, that by Assumption 1 functions $x \mapsto \pi_k(x)$ are Lipschitz for all $k = -K \dots K$. Therefore we can select $C = \max\{C_{-K} \dots C_K\}$ s.t. $|\pi_k(x) - \pi_k(y)| \leq C|x - y|$ for all x, y in the state-space of the chain. Hence, we arrive at the following inequality:

$$
|P_x f - P_y f| \leq \epsilon' + (2K + 1)C|x - y| .
$$

   Next, consider two-step transition probabilities. Using similar arguments we have:

$$
|P_x^2 f - P_y^2 f| = |P_x(P_{x'} f) - P_y(P_{y'} f)|
$$

$$\leq \sum_{k=-K}^{K} |P_{x'}f - P_{y'}f|\pi_k(x) + \sum_{k=-K}^{K} |\pi_k(x) - \pi_k(y)|\,,$$

where $x' = \gamma_0 + \gamma_1 x - \gamma_2 \mathbb{E}(\Delta p^2|x) + \gamma_2 k^2$ and analogously for y'. Then $|x' - y'| = \Big|\gamma_1(x-y) - \gamma_2(\mathbb{E}(\Delta p^2|x) - \mathbb{E}(\Delta p^2|y))\Big| \leq \phi|x-y|$ by Assumption 2, where $\phi < 1$. Hence we have:

$$|P_x^2 f - P_y^2 f| \leq \epsilon' + (2K+1)C\phi|x-y| + (2K+1)C|x-y|\,.$$

By induction,

$$\begin{aligned}|P_x^n f - P_y^n f| &\leq \epsilon' + (2K+1)C|x-y|\sum_{j=0}^{n-1}\phi^j \\ &\leq \epsilon' + \frac{(2K+1)C}{1-\phi}|x-y| \\ &\leq \epsilon' + \frac{(2K+1)C}{1-\phi}\epsilon_2 \leq \epsilon_1\,.\end{aligned}$$

Hence, collection of Markov transition kernels $\{P^n f : n \geq 1\}$ is equicontinuous for IV-GARCH(1,1) model.

3. Lastly, we show that the point $\{\underline{\sigma}^2\}$ is a reachable state of the chain $\{\sigma_i^2\}_{i\geq 1}$. It is enough to show that for any open $O \in \mathfrak{B}(\mathcal{X})$ containing $\{\underline{\sigma}^2\}$ there exists $1 \leq n < \infty$ s.t. $P^n(x, O) > 0$ for any starting value x in the state-space of the chain.

   From equation (5) we see that σ_i^2 can be written as:

$$\sigma_i^2 = \gamma_0 \sum_{j=0}^{i-1}\gamma_1^i + \gamma_1^i\sigma_0^2 + \gamma_2\sum_{j=0}^{i-1}\gamma_1^{i-1-j}\Delta p_j^2 - \gamma_2\sum_{j=0}^{i-1}\gamma_1^{i-1-j}\mathbb{E}(\Delta p_j^2|\sigma_j^2)\,.$$

   Consider the case when $\{\Delta p_i^2\}_{i\geq 1}$ is a sequence of zero price innovations, where each zero price innovation has probability $\pi_0(\sigma_i^2)$, which by Assumption 1 is strictly greater than zero for all σ_i^2 in the state-space of the chain. By Proposition 1 the limit of σ_i^2 is given by:

$$\begin{aligned}\lim_{i\to\infty}\sigma_i^2 &= \lim_{i\to\infty}\gamma_0\sum_{j=0}^{i-1}\gamma_1^i + \lim_{i\to\infty}\gamma_1^i\sigma_0^2 - \lim_{i\to\infty}\gamma_2\sum_{j=0}^{i-1}\gamma_1^{i-1-j}\mathbb{E}(\Delta p_j^2|\sigma_j^2) \\ &= \underline{\sigma}^2\,,\end{aligned}$$

   and by definition of the limit there exist $1 \leq n < \infty$ s.t. σ_i^2 is arbitrary close to $\{\underline{\sigma}^2\}$ with probability $\prod_{i=0}^{n}\pi_0(\sigma_i^2) > 0$. □

2.5 IV-FIARCH model for heteroscedastic discrete Δp_i

The long-memory version of the model in equation (3), referred to as IV-FIARCH(∞) model, is defined by:

$$\begin{aligned}\Delta p_i &\sim \pi_{-K}(\mu_i, \sigma_i^2; \boldsymbol{\alpha}), \ldots \pi_K(\mu_i, \sigma_i^2; \boldsymbol{\alpha}) \\ \mu_i &= \boldsymbol{x}_i'\boldsymbol{\beta} \\ \sigma_i^2 &= \gamma_{0i} + \gamma_2(1-L)^{-d}\Big(\Delta p_{i-1}^2 - \mathbb{E}(\Delta p_{i-1}^2|\sigma_{i-1}^2)\Big)\,,\end{aligned} \tag{8}$$

where d is the coefficient of fractional integration; see Baillie (1996) for a recent survey of the literature on fractional integration in econometric. In this model $\mathbb{E}(\sigma_i^2) = \mathbb{E}(\gamma_{0i})$.

First, we give sufficient condition for non-negativity of the volatility process in IV-FIARCH(∞) model:

Lemma 2 *Volatility process of IV-FIARCH(∞) model defined in equation (8) is non-negative if* $\sigma^2 \mapsto \sigma^2 - \gamma_2 \mathbb{E}(\Delta p^2 | \sigma^2) \geq 0$.

PROOF Similarly to the case of IV-GARCH(1,1) model, we write noise-free skeleton of the process $\{\sigma_i^2\}_{i\geq 1}$ in (8) as follows:

$$\sigma_i^2 = \gamma_0 - \gamma_2 \sum_{j=1}^{i} \psi_{j-1} \mathbb{E}(\Delta p_{i-j}^2 | \sigma_{i-j}^2) , \tag{9}$$

where coefficients ψ_j come from expansion of the polynomial $(1 - L)^{-d}$; see Hosking (1981). Observe the following recursion:

$$\begin{aligned}
\sigma_1^2 &= \gamma_0 - \gamma_2 \mathbb{E}(\Delta p_0^2 | \sigma_0^2) \\
\sigma_2^2 &= \gamma_0 - \gamma_2 \mathbb{E}(\Delta p_0^2 | \sigma_0^2) + \gamma_2 \mathbb{E}(\Delta p_0^2 | \sigma_0^2) - \gamma_2 \psi_1 \mathbb{E}(\Delta p_0^2 | \sigma_0^2) - \gamma_2 \mathbb{E}(\Delta p_1^2 | \sigma_1^2) \\
&= \sigma_1^2 - \gamma_2 \mathbb{E}(\Delta p_1^2 | \sigma_1^2) + \gamma_2 (1 - \psi_1) \mathbb{E}(\Delta p_0^2 | \sigma_0^2) \\
\sigma_3^2 &= \ldots = \sigma_2^2 - \gamma_2 \mathbb{E}(\Delta p_2^2 | \sigma_2^2) + \gamma_2 (1 - \psi_1) \mathbb{E}(\Delta p_1^2 | \sigma_1^2) + \gamma_2 (\psi_1 - \psi_2) \mathbb{E}(\Delta p_0^2 | \sigma_0^2) \\
&\vdots
\end{aligned}$$

Assume $\sigma_1^2 > 0$. By induction we have:

$$\sigma_i^2 = \gamma_0 - \mathbb{E}(\Delta p_{i-1}^2 | \sigma_{i-1}^2) + \gamma_2 \sum_{j=2}^{i} (\psi_{j-2} - \psi_{j-1}) \mathbb{E}(\Delta p_{i-j}^2 | \sigma_{i-j}^2) ,$$

from where sufficiency of $\sigma^2 \mapsto \sigma^2 - \gamma_2 \mathbb{E}(\Delta p^2 | \sigma^2) \geq 0$ follows by positivity of $(\psi_{j-1} - \psi_j)$ for $j \geq 1$. □

An immediate consequence of Lemma 2 is that $\lim_{\sigma^2 \to 0} \mathbb{E}(\Delta p^2 | \sigma^2) = 0$. This implies $\delta = 0$ in Assumption 1, part 2, in the IV-FIARCH(∞) model.

In the following proposition we establish deterministic stability of the volatility process in the IV-FIARCH(∞) model:

Proposition 2 *Under conditions of Lemma 2, deterministic part of the volatility process in IV-FIARCH(∞) model defined by* $\sigma_i^2 = \gamma_0 - \gamma_2 \sum_{j=1}^{i} \psi_{j-1} \mathbb{E}(\Delta p_{i-j}^2 | \sigma_{i-j}^2)$ *is globally stable and has a unique limit point* $\underline{\sigma}^2 = 0$.

PROOF From Lemma 2 follows that, starting from any $\sigma_0^2 \in \mathcal{R}_+$, the sequence $\{\sigma_i^2\}_{i\geq 1}$ from the noise-free skeleton (9) of the volatility process in IV-FIARCH(∞) model is bounded below by zero. At the same time, equation (9) implies that the sequence $\{\sigma_i^2\}_{i\geq 1}$ is monotonically decreasing. Hence, the result follows from the convergence theorem for monotonic bounded sequences; see Theorem 3.14 in Rudin (1976). □

So far we introduced two simple dynamic heteroscedasticity models for high-frequency price changes. As seen from equations (3) and (8), IV-GARCH(1,1) and IV-FIARCH(∞) models have relatively restricted short- and long-memory dynamics. The work is being done to extend the models to have richer dynamics, including short-memory part in the IV-FIARCH model.

2.6 Estimation and diagnostic procedures for IV-GARCH and IV-FIARCH models

Statistical inference in IV-GARCH and IV-FIARCH models is based on the straightforward maximum-likelihood procedure. Given the set of observed data $\{\Delta p_i, \boldsymbol{x}_i, \boldsymbol{z}_i\}_{i=1}^N$ the log-likelihood function is given by:

$$\log L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta} | \{\Delta p_i, \boldsymbol{x}_i, \boldsymbol{z}_i\}_{i=1}^N) = \sum_{i=1}^{N} \sum_{k=-K}^{K} \mathbf{1}_{\Delta p_i = k} \cdot \pi_k \, .$$

In IV-FIARCH model the parameter d replaces γ_1 in the log-likelihood function. In the empirical part of the paper in section 4 we employ numerical procedures to calculate the gradient and the Hessian matrix of the log-likelihood function. We use BHHH method of Berndt, Hall, Hall and Hausman (1974) for numerical maximization of the log-likelihood functions. Standard errors of the parameter estimates are computed from the diagonal elements of the inverted negative Hessian matrix at the point of maximum of the log-likelihood function.

Diagnostics in IV-GARCH and IV-FIARCH models can be based on the generalized residuals calculated as follows:

$$\hat{u}_i = \frac{\Delta p_i - \mathbb{E}(\Delta p_i | \sigma_i^2)}{\sqrt{\mathbb{V}(\Delta p_i | \sigma_i^2)}} \, . \tag{10}$$

By construction, under the maintained hypothesis of Δp_i generated from IV-GARCH or IV-FIARCH model, the sequence $\{\hat{u}_i\}_{i=1}^N$ will be homoscedastic and uncorrelated. A battery of standard test procedures can be applied to $\{\hat{u}_i\}_{i=1}^N$ to test for this assumptions in the real-world data.

3 Data and descriptive statistics

In this section we give an overview of the high-frequency transaction data used in the empirical part of the paper in section 4. We present simple descriptive statics of the dataset, discuss evidence of long-range dependence in the volatility of high-frequency price changes and its dependence on the tails of price changes distribution.

In section 4 we apply IV-GARCH and IV-FIARCH models to the high-frequency IBM trades data. This dataset has been used in the series of recent papers by Engle and Russel (1998), Russel and Engle (1998) and Engle (2000). The dataset originates from the TORQ database and covers trades of the IBM stock on the weekdays during the three-month period from November, 1990 through January, 1991. There are 60328 observations in the sample, and besides the date and the timestamp the dataset also includes information on the traded volume, transaction price and the bid-ask price of the stock at the time of the trade.

From this dataset we extract transaction prices and create price changes data Δp_i by taking their first difference. The prices and price changes are discrete due to the institutional structure of the NYSE where the stock is traded. Δp_i series is the multiple of $\$\frac{1}{8}$, and we delete a few "keypunch" errors where this has not been the case. In addition to the high-frequency price changes, we calculate inter-trade durations and construct trading hour indicator dummies for picking up intraday seasonal effects. We also create buyer-seller indicators by comparing the current transaction price with the mid-quote price coming at least 5 seconds before the transaction; see Campbell, Lo and MacKinlay (1997) pp. 136-137 for the discussion of this methodology.

It has become a common practice in the literature to filter out observations that have zero inter-trade durations and zero price changes; see Russel and Engle (1998) and Jasiak (1999) among others. This adjustment is believed to help reducing the influence of so-called splitted trades, whereby larger orders are divided into a number of smaller ones traded at the same price. Therefore, the filtered price changes can be attributed to the unique transactions leading to the price movements due to the new information arrivals and/or position adjustments/liquidity considerations. In the empirical application of IV-GARCH and IV-FIARCH models in section 4 we also follow this procedure for the IBM data.

Table 1: Probabilities of Δp_i.

s_k	$\Pr(\Delta p_i = s_k)$
$\leq$ -0.50000	0.0042211
-0.37500	0.0027775
-0.25000	0.014217
-0.12500	0.16002
0.0000	0.63633
0.12500	0.16016
0.25000	0.014947
0.37500	0.0031064
$\geq$ 0.50000	0.0042211

Distribution of the resulting Δp_i series is shown on figure 4. Table 1 shows the state probabilities of high-frequency Δp_i, where we censor price changes higher than $\$\frac{1}{2}$ in absolute value. Although the distribution of Δp_i series on figure 4 is notably symmetric around its middle state, table 1 show that the right half of the distribution has slightly higher probability mass reflecting the upward drift of the IBM stock price during the sample period. Nevertheless, the assumption of the symmetry of unconditional distribution of price changes in IV-GARCH and IV-FIARCH models made in section 2 seems warranted in the view of the evidence in table 1 and on figure 4. Note that the parameter μ_i in IV-GARCH and IV-FIARCH models is designed to pick up short-term fluctuations in the first moment of high-frequency price changes, and is likely to have only limited success in describing its long term trend.

FIGURE 4 IS ABOUT HERE

Descriptive statistics of the Δp_i series together with its transformations is given in table 2. After all data adjustments there are 54725 observations in the sample. It is seen that the sample average of the series is statistically indistinguishable from zero, and that small variance of Δp_i reflects the dominating probability of zero price movement in the data. Symmetry of the distribution of Δp_i is confirmed by the statistically insignificant skewness statistic. Portmanteau statistics presented in the table clearly indicate the presence of the dynamic structure in both the first and second moments of high-frequency price changes. This is further confirmed by the correlograms of Δp_i series on figure 5[9]. It is seen that, while there is a significant negative first-order autocorrelation in Δp_i series, all higher-order autocorrelations are statistically insignificant. This patterns corresponds very well to the bid-ask bounce model of Roll (1984). In the empirical part of the paper in section 4 we include lags of buyer-seller indicator in our specification of μ_i to capture this effect.

Table 2: Sample statistics of high-frequency price changes.

	Δp_i	$\lvert\Delta p_i\rvert$	Δp_i^2
Mean	0.000342622	0.0558748*	0.0140304*
Variance	0.0140303	0.0109084	0.0244471
Skewness	-0.02079	8.3517*	61.4053*
Kurtosis	125.193*	187.434*	4555.33*
Q(500)	12773.6*	42829*	32837.3*
Maximum	3.625	3.625	13.1406
Minimum	-3.625	0	0
No. of obs.:	54725	54725	54725

Notes: Δp_i denotes high-frequency price change for IBM transaction data. Tick size equals to $\$\frac{1}{8}$. Skewness and kurtosis statistics and their standard errors are according to Jarque and Bera (1987). Q(500) denotes Ljung-Box statistic with autocorrelations up to 500 lags; see Ljung and Box (1978). Star near a test statistic indicates significance on 5% level for the appropriate distribution.

Lower panel of figure 5 reveals substantial degree of persistence in the second moment of high-frequency price changes in IBM trades data as given by 500 lags of significant autocorrelations in $|\Delta p_i|$ series. Similar pattern in the irregularly spaced stock market data has been documented earlier in Rydberg and Shephard (2000), while Andersen and Bollerslev (1997a, 1997b, 1998) document long-range dependence in the volatility of five-minute foreign exchange returns.

FIGURE 5 IS ABOUT HERE

[9]Under i.i.d. normality assumption 95% confidence band for estimated autocorrelations is given by $\pm\frac{1.96}{\sqrt{T}}$. Confidence intervals for correlograms shown on figure 5 are given by ± 0.0084 for both Δp_i and $|\Delta p_i|$ series.

Number of parameters in IV-GARCH and IV-FIARCH models in section 2 depends on the number of support points in the distribution of high-frequency price changes. As shown in table 1, the states of Δp_i far in the tails of the distributions have quite small probability. Therefore, in section 4 of the paper we estimate IV-GARCH and IV-FIARCH models with reduced number of states of Δp_i, saving on the estimation time by cutting the parameters that are likely to be estimated inefficiently. However, there has been little evidence in the literature on the effects of the censoring of large price changes in high-frequency data on the dynamics of the volatility of Δp_i series. Figure 6 depicts correlograms of $|\Delta p_i|$ series censored to 7, 5 and 3 states[10]. It is seen that the very extreme states of the price changes distribution do not play a major role in the volatility dynamics of the series. When Δp_i is censored down to 3 states from the initial number of 37 states — equivalent of loosing less than 5% of the information compared to the initial distribution — the autocorrelations of absolute price changes become remarkably low, becoming statistically insignificant already after 50 lags. With 7 states in price changes distribution — still much lower than the number of states in the uncensored data — the dynamics of $|\Delta p_i|$ closely resembles that of the original series on figure 5.

FIGURE 6 IS ABOUT HERE

4 Fitting IV-GARCH and IV-FIARCH models to IBM transaction data

In this section we report estimation results for IV-GARCH and IV-FIARCH models using IBM transaction dataset introduced in section 3. We study the influence of several explanatory variables on the conditional volatility part of IV-GARCH and IV-FIARCH models, but our main goal is to gauge the overall success of these two models in explaining volatility dynamics of discrete high-frequency data. We present model diagnostics based on the generalized residuals from equation (10) on page 22 to assess the quality of the fit[11].

In the application of IV-GARCH and IV-FIARCH models to IBM data we censor high-frequency price changes to 7 states. This allows us save on the number of estimated parameters, but at the same time preserves essential dynamic structure of the volatility in the data; see section 3. With 7 support points of the estimated distribution of Δp_i the dimensionality of $\boldsymbol{\alpha}$ parameter equals to 6, out of which 2 parameters are free.

First, we discuss estimation results for the IV-GARCH model reported in the left half of table 3. Apart from the constant term, parameter μ_i of IV-GARCH model includes the following explanatory variables:

- Ibs_{-1} is lagged buyer-seller indicator constructed as outlined in section 3. We include one lag of this variable to pick up the bid-ask bounce in the first moment of Δp_i series.

[10] The confidence band for autocorrelations on figure 6 is given by ± 0.0084 for all three graphs.

[11] Graphical illustrations in the paper and model estimations in this section are made in Ox version 2.20 for Linux 2.2.17 — an excellent programming language for econometricians; see Doornik (1998).

- Δp_{-1} is included to account for possible autoregressive dynamics of the first moment of Δp_i series. Hausman, Lo and MacKinlay (1992) found the lags of endogenous variable to be significant in the conditional mean of their ordered probit model.

- *Durat* is the inter-trade duration variable for the current observation. We treat this variable as exogenous with respect to the price changes in all models in this section, although this assumption may not be entirely realistic; see Ghysels and Jasiak (1997) for the evidence of why this may not be so.

As was mentioned in section 2, all variables in the conditional mean part of IV-GARCH model, with the exception of the constant term, should be scaled to have zero means. This is done for all exogenous variables entering μ_i in IV-GARCH model.

Separately from μ_i, we study the effect of several explanatory variables on the conditional volatility part of IV-GARCH model:

- $Negdp_{-1}$ is the indicator variable showing occurrence of the previous negative price change in the series. This variable is included to study possible leverage effect of Nelson (1991) in the high-frequency data. Rydberg and Shephard (1999) found this variable to be significant in their ADS decomposition model.

- *Durat* is the same variable that appears in the conditional mean part of the model. By including inter-trade duration variable two times we study the effect of trading intensity separately on the first and second moments of Δp_i. Hausman, Lo and MacKinlay (1992) found this variable marginally significant in static conditional volatility part of their model.

- Trade hour dummies pick up possible intraday seasonality in the second moment of high-frequency price changes. Intraday seasonal patterns in the volatility of high-frequency data are widely reported for the regularly spaced data; see Andersen, Bollerslev and Cai (2000) for the recent evidence.

Signs of the coefficients for the variables listed above were not restricted during the estimation procedure. Therefore, the direction of the influence of included explanatory variables on the second moment of high-frequency price changes will coincide with the signs of the estimated parameters.

As seen from table 3, γ_1 parameter of the conditional volatility process is highly significant, but is firmly below unity. With the value of this parameter estimated at 0.93478, the half-live of a unit shock to the conditional variance process is given by only 10 periods. This seems to contradict the evidence of the long-range dependence in the volatility of Δp_i series presented in section 3. Generalized residuals diagnostics presented in the bottom half of table 3 also hints to the unexplained dynamics left in both the first and second moments of Δp_i series. However, graphical examination of the correlograms of generalized residuals and absolute

Table 3: Results of modeling high-frequency IBM trades series.

	IV-GARCH		IV-FIARCH	
		***α** parameters*		
α_1	2.5553	0.022118	2.5555	0.039110
α_2	3.4401	0.041139	3.3469	0.064460
		Conditional mean parameters		
Const	0.0043945	0.0068764	0.0086375	0.013443
Ibs_{-1}	-0.30770	0.0084919	-0.32839	0.016626
Δp_{-1}	-5.3752	0.091196	-5.4225	0.16522
Durat	-0.055171	0.0095503	-0.10364	0.030933
		Conditional variance parameters		
Const	0.030987	0.0037965	1.5471	0.44586
$Negdp_{-1}$	-0.024142	0.0071386	0.083911	0.021269
Durat	-0.0017309	0.0015081	-0.032359	0.021301
9-10h	-0.0011467	0.0016714	-0.21259	0.047269
11-12h	-0.0016986	0.0016620	-0.16512	0.063266
13-14h	-0.0020903	0.0017012	-0.17244	0.064233
γ_1	0.93478	0.0086118	—	—
γ_2	0.069761	0.0060288	0.10259	0.011412
d	—	—	0.60528	0.033268
Skewness	-0.153967*		-0.132489*	
Kurtosis	5.55161*		5.35271*	
Q(500)	1777.92*		1081.83*	
$Q^2(500)$	729.472*		594.943*	
Log lik.	-27929.287		-8619.459	
No of obs.	54725		22000	

Notes: Parameters of the models are denoted as detailed in the text. Asymptotic maximum-likelihood standard errors are given in the parenthesis. Generalized residuals calculated according to equation 10. Skewness and kurtosis statistics and their standard errors are according to Jarque and Bera (1987). Q(500) denotes Ljung-Box statistic of the generalized residuals and $Q^2(500)$ of squared generalized residuals with autocorrelations up to 500 lags; see Ljung and Box (1978). Star near a test statistic indicates significance on 5% level for the appropriate distribution.

generalized residuals on figure 7[12]. reveals dramatic reduction of the volatility dynamics of the residuals compared to the original series. In fact, as suggested by the lower right panel of figure 7, the significance of LB statistic of squared residuals stems from the unexplained fist-order autocorrelation in the volatility of Δp_i series, indicating possible omission of explanatory variables or the need for extra MA dynamics in σ_i^2. Most importantly, however, IV-GARCH model seems to do a good job in picking up the conditional heteroscedasticity of high-frequency price changes.

FIGURE 7 IS ABOUT HERE

Most coefficients of other explanatory variables in IV-GARCH model in table 3 have expected signs. Buyer-seller indicator Ibs_{-1} have expected negative influence on the conditional mean part of the model, indicating increased probability of negative price change when the transaction is seller-initiated. However, as seen from the lower left panel of figure 7, generalized residuals still retain significant fist order negative autocorrelation. This may signal the failure of Ibs_{-1} variable to correctly classify all transactions in the dataset into either initiated by the buyer or seller[13]. *Durat* variables is only significant in the conditional mean part, and is also indicating right skew of the distribution of Δp_i for longer inter-trade durations. As in the static model of Hausman, Lo and MacKinlay (1992), inter-trade durations have no significant influence in the conditional volatility part of the model. Somewhat surprising result shown in table 3 is the effect of the previous negative price change on the conditional volatility of Δp_i. In contrast to the leverage hypothesis of Nelson (1991), conditional volatility of high-frequency Δp_i in IBM data becomes lower after the preceding stock price drop, although the estimated coefficient of $Negdp_{-1}$ is not very significant for the given sample size. Effect of intraday seasonality dummies on σ_i^2 is also insignificant.

Estimation results for IV-FIARCH model are given in the right half of table 3[14]. We use the same set of exogenous explanatory variables as in the IV-GARCH model. The estimate of the coefficient of fractional integration is above 0.5 and highly statistically significant. Figure 8[15]. shows high persistence of the volatility of Δp_i in the estimation subsample as well as dramatic reduction of this persistence in the estimated generalized residuals, although similarly to the IV-GARCH model some significant low-order autocorrelation is still present. Effects of the included explanatory variables remain the same, except for the $Negdp_{-1}$ in the conditional variance part of the model. Even though it now supports the leverage effect of Nelson (1991), the estimated standard error of the coefficient is relatively large for the given sample size.

FIGURE 8 IS ABOUT HERE

[12] The confidence band for autocorrelations on figure 7 is given by ± 0.0084 for all three graphs.

[13] In is also necessary to note here that the methodology of creating the Ibs_{-1} variable does not allow for the full data classification. In our dataset around 21% of observations are left unclassified, which may also contribute to the remaining first-order autocorrelation in the generalized residuals.

[14] Please note that the model is estimated on the reduced sample due to computational time considerations.

[15] The confidence band for autocorrelations on figure 8 is given by ± 0.0132 for all three graphs.

Surprisingly, both IV-GARCH and IV-FIARCH models are doing very much alike in terms of the residuals diagnostics shown on figures 7 and 8, even though the AR(1) structure of the volatility process of the IV-GARCH model is not able to pick up high persistence exhibited by the data. Relatively low point estimate of γ_1 in the model reported in table 3 seems to reaffirm the analogous results documented in Ghysels and Jasiak (1997) and Hasbrouck (1999); see subsection 1.4. In summary, the findings call for further investigation into potential infrequent changes in the volatility regimes on the market that can create statistical illusion of the long-range dependence in the data.

Conclusions

In this paper we present a class of model for time-series of discrete high-frequency price changes. In contrast to the ACM model of Russel and Engle (1998) and the ADS decomposition model of Rydberg and Shephard (1999), our models have separate set of parameters linked to the first two moments of the conditional distribution of discrete price changes. We borrow the idea of Hausman, Lo and MacKinlay (1992) and specify discrete distribution of Δp_i similarly to the well-known ordered logit model. But unlike the latter study, our models are cast entirely in terms of discrete random variables, including procedures for model diagnostics. We introduce IV-GARCH and IV-FIARCH models for heteroscedastic sequences of discrete price changes, where volatility parameter has the dynamic structure resembling the one in GARCH models of Bollerslev (1986).

Separate sets of parameters for the moments of discrete price changes allow us to isolate effects of exogenous variables on conditional mean and conditional variance of Δp_i. We present application of IV-GARCH(1,1) and IV-FIARCH(∞) models to high-frequency IBM trades data, where we study effects of inter-trade durations, buyer-seller indicator and previous negative price changes on the moments of high-frequency price changes.

We find that both IV-GARCH and IV-FIARCH models explain dynamic heteroscedasticity of Δp_i series quite well. In particular, both models succeed in explaining most of the long-range dependence observed in absolute price changes series in the data, although some low order dependence remains. Unexpectedly, short-memory IV-GARCH(1,1) model fits the dataset at least as good as the IV-FIARCH(∞) model in terms of residuals diagnostics. This may indicate that observed long-range dependence in $|\Delta p_i|$ comes from the infrequent changes in the volatility regimes of the stock market, rather than from the generic long-memory structure in the second moment.

Current research efforts in the literature are directed to joint modeling of price changes and durations in high-frequency financial data; see Gerhard and Pohlmeier (2000). In this paper we estimate conditional model of price changes volatility and find an insignificant effect of the immediately preceding inter-trade duration on σ_i^2 parameter both in IV-GARCH and IV-FIARCH models. This finding is surprising and calls for further investigation of the interaction between the two variables in the joint model. Combination of the ACD model of Engle and Russel (1998) and IV-GARCH model introduced in this paper may provide a useful tool for such analysis.

The interrelations between possible volatility regimes and long-range dependence in the second moment of high-frequency price changes is another research issue raised by the findings in this paper. The literature on potential links between structural breaks and long memory in the volatility of financial data is numerous (see Hamilton and Susmel (1994), Lamoureux and Lastrapes (1990) and Liu (2000) among others), but is mostly limited to lower frequency financial data. IV-FIARCH model offers an opportunity to study this issue in high-frequency datasets.

References

Andersen, Torben G. and Tim Bollerslev (1998a) Deutsche Mark-Dollar volatility: intraday activity patterns, macroeconomic announcements and longer run dependencies. *Journal of Finance*, vol. 53, pp. 219-265.

Andersen, Torben G. and Tim Bollerslev (1997a) Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, vol. 4, pp. 115-158.

Andersen, Torben G. and Tim Bollerslev (1997b) Heterogeneous information arrivals and returns volatility dynamics: uncovering the long-run in high frequency returns. *Journal of Finance*, vol. 52, pp. 975-1005.

Andersen, Torben G., Tim Bollerslev and Jun Cai (2000) Intraday and interday volatility in the Japanese stock market. *Journal of International Financial Markets, Institutions and Money*, vol. 10, pp. 107-130.

Baillie, Richard T. (1996) Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, vol. 73, pp. 5-59.

Ball, C. (1988) Estimation bias induced by discrete security prices. *Journal of Finance*, vol. 43, pp. 841-865.

Berndt, E., B. Hall, R. Hall and J. Hausman (1974) Estimation and inference in non-linear structural models. *Annals of Economic and Social Measurement*, vol. 3, pp. 653-665.

Bollerslev, Tim (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, vol. 31, pp. 307-327.

Bollerslev, Tim, Ray F. Chou and Kenneth F. Kroner (1992) ARCH modeling in finance. *Journal of Econometrics*, vol. 52, pp. 5-59.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay (1997) *The econometrics of financial markets.* Princeton, New Jersey: Princeton University Press.

Cox, D. R. (1981) Statistical analysis of time series: some recent developments. *Scandinavian Journal of Statistics*, vol. 8, pp. 93-115.

Crack, Timothy Falcon and Olivier Ledoit (1996) Robust structure without predictability: the "compass rose" pattern of the stock market. *Journal of Finance*, vol. 51, no. 2, pp. 751-762.

Davis, A. Richard, Tina Hviid Rydberg, Neil Shephard and Sarah B. Streett (2001) The CBIN model for counts: testing for common features in the speed of trading, quote changes, limit and market order arrivals. Mimeo.

Doornik, Jurgen A. (1998) *Object-Oriented Matrix Programming using Ox 2.0.* London: Timberlake Consultants Ltd and Oxford: `www.nuff.ox.ac.uk/Users/~Doornik`.

Drost, Feike C. and Theo E. Nijman (1993) Temporal aggregation of GARCH processes. *Econometrica*, vol. 61, no. 4, pp. 909-927.

Dufour, Alfonso and Robert F. Engle (2000) Time and the price impact of a trade. *Journal of Finance*, vol. 55, no. 6, pp. 2467-2498.

Engle, Robert F. (2000) The econometrics of ultra-high-frequency data. *Econometrica*, vol. 68, no. 1, pp. 1-22.

Engle, Robert F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, vol. 50, pp. 987-1008.

Engle, Robert F. and Jeffrey R. Russel (1998) Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, vol. 66, pp. 1127-1162.

Feller, William (1968) *An Introduction to Probability Theory and its Applications*, 3rd edition, Wiley.

Gerhard, Frank and Winfried Pohlmeier (2000) On the Simultaneity of Components of the Transaction Process. University of Konstanz. Mimeo.

Ghysels, Eric (2000) Some econometric recipes for high-frequency data cooking. *Journal of Business and Economic Statistics*, vol. 18, no. 2, pp. 154-163.

Ghysels, Eric and Joanna Jasiak (1997) GARCH for irregularly spaced financial data: the ACD-GARCH model. Mimeo.

Gottlieb, Gary and Avner Kalay (1985) Implications of the discreteness of observed stock prices. *Journal of Finance*, vol. 40, pp. 135-153.

Glynn, Peter and Sean Meyn (1997) Tightness for non-irreducible Markov chains. Mimeo.

Hamilton, James D. and Raul Susmel (1994) Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, vol. 64, pp. 307-333.

Hasbrouck, Joel (1999) The dynamics of discrete bid and ask quotes. *Journal of Finance*, vol. 54, no. 6,pp. 2109-2142.

Hausman, Jerry A., Andrew W. Lo and A. Craig MacKinlay (1992) An ordered probit analysis of transaction stock prices", *Journal of Financial Economics*, vol. 31, pp. 319-379.

Hautsch, Nikolas and Winfried Pohlmeier (2001) Econometric analysis of financial transaction data: pitfalls and opportunities. University of Konstanz. Mimeo.

Hosking, Jonathan R.M. (1981) Fractional differencing. *Biometrika*, vol. 68, pp. 165-76.

Jarque, C. M. and A. K. Bera (1987) A test for normality of observations and regression residuals. *International Statistical Review*, vol. 55, pp. 163-172.

Jasiak, Joanna (1999) Persistence in intertrade durations. Mimeo.

Johnston, Norman L. and Samuel Kotz (1969) *Discrete distributions*, Houghton Mifflin Company, Boston.

Lamoureux, Christopher G. and William D. Lastrapes (1990) Persistence in variance, structural change and the GARCH model. *Journal of Business and Economic Statistics*, vol. 8, pp. 225-234.

Liu, Ming (2000) Modeling long memory in stock market volatility. *Journal of Econometrics*, vol. 99, pp. 139-171.

Ljung, G. M. and G. P. E. Box (1978) On a measure of lack of fit in time series models. *Biometrika*, vol. 66, pp. 66-72

Meyn, Sean and Richard L. Tweedie (1993) *Markov chains and stochastic stability*, Springer-Verlag, London.

McFadden, Daniel L. (1984) *Econometric analysis of qualitative response models.* In Handbook of Econometrics, vol. II, Elsevier Science, North-Holland.

Nelson, Daniel B. (1991) Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, vol. 59, pp. 347-370.

Roll, R. (1984) A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, vol. 39, pp. 1127-1140.

Rudin, Walter (1976) *Principles of mathematical analysis.* Third Edition, McGraw-Hill International.

Russell, Jeffrey R. and Robert F. Engle (1998b) Econometric analysis of discrete-valued irregularly-spaced financial transactions data using a new autoregressive conditional multinomial model. Mimeo.

Rydberg, Tina Hviid and Neil Shephard (2000) A modelling framework for the prices and times of trades made on the New York stock exchange. Mimeo.

Rydberg, Tina Hviid and Neil Shephard (1999) Dynamics of trade-by-trade price movements: decomposition and models. Mimeo.

Szpiro, George G. (1998) Tick size, the compass rose and market nanostructure. *Journal of Banking and Finance*, vol. 22, pp. 1559-1569.

Tong, Howell (1990) *Non-linear time series. A dynamic system approach.* New York, Oxford University Press.

Tweedie, Richard L. (1998) Markov chains: structure and applications. Mimeo.

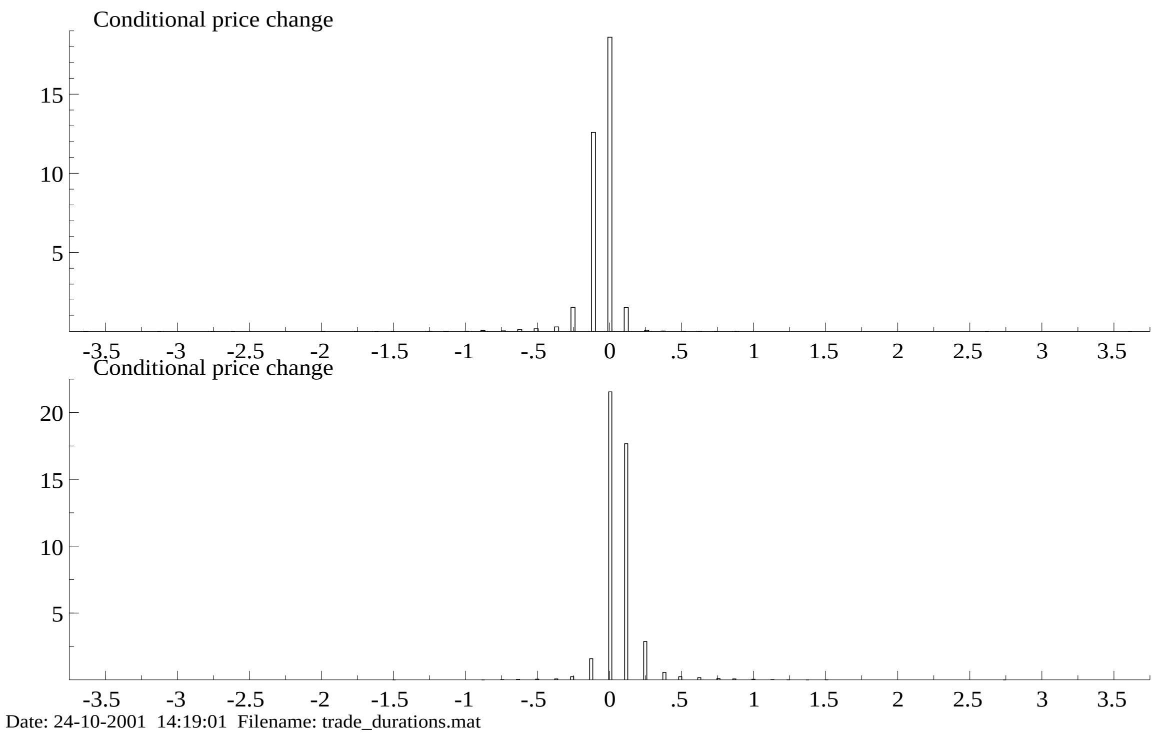

Figure 1: Distribution of Δp_i conditional on $\Delta p_{i-1} \geq 0$ (upper panel) and on $\Delta p_{i-1} \leq 0$ (lower panel) in IBM transaction data. Tick size equals to $\$\frac{1}{8}$.

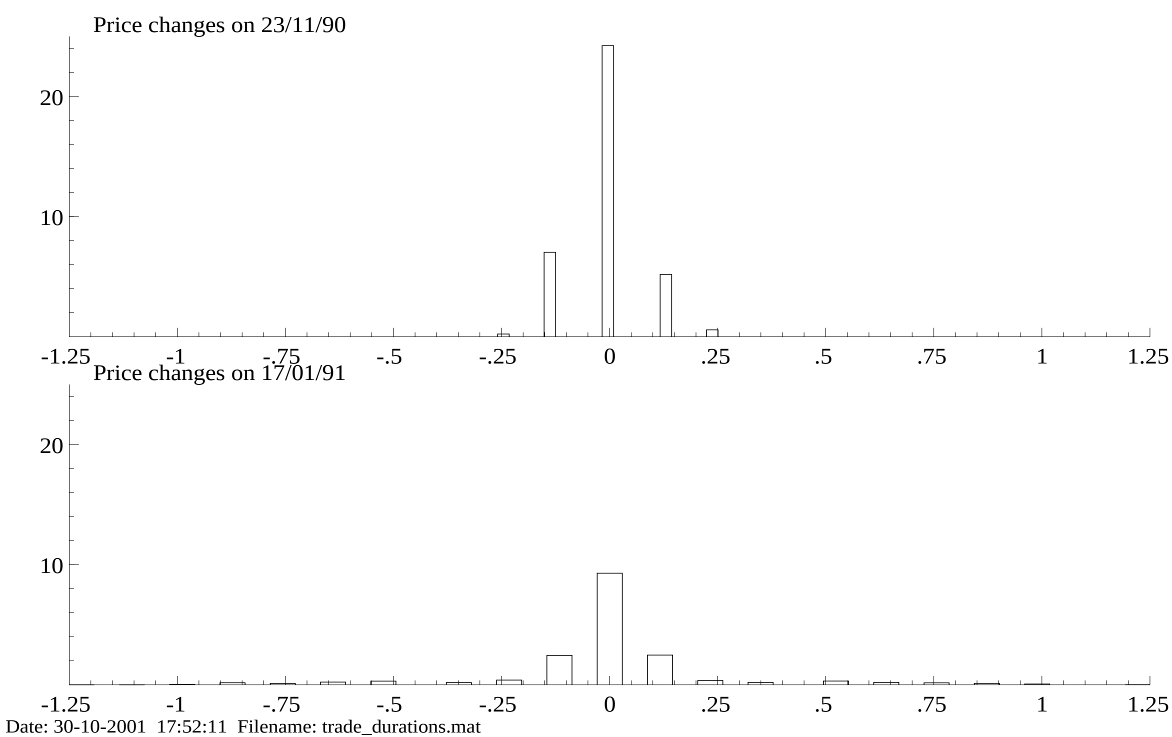

Figure 2: Marginal distributions of Δp_i on 23^{rd} of November 1990 (upper panel) and 17^{th} of January 1991 (lower panel) in IBM transaction data. Tick size equals to $\$\frac{1}{8}$.

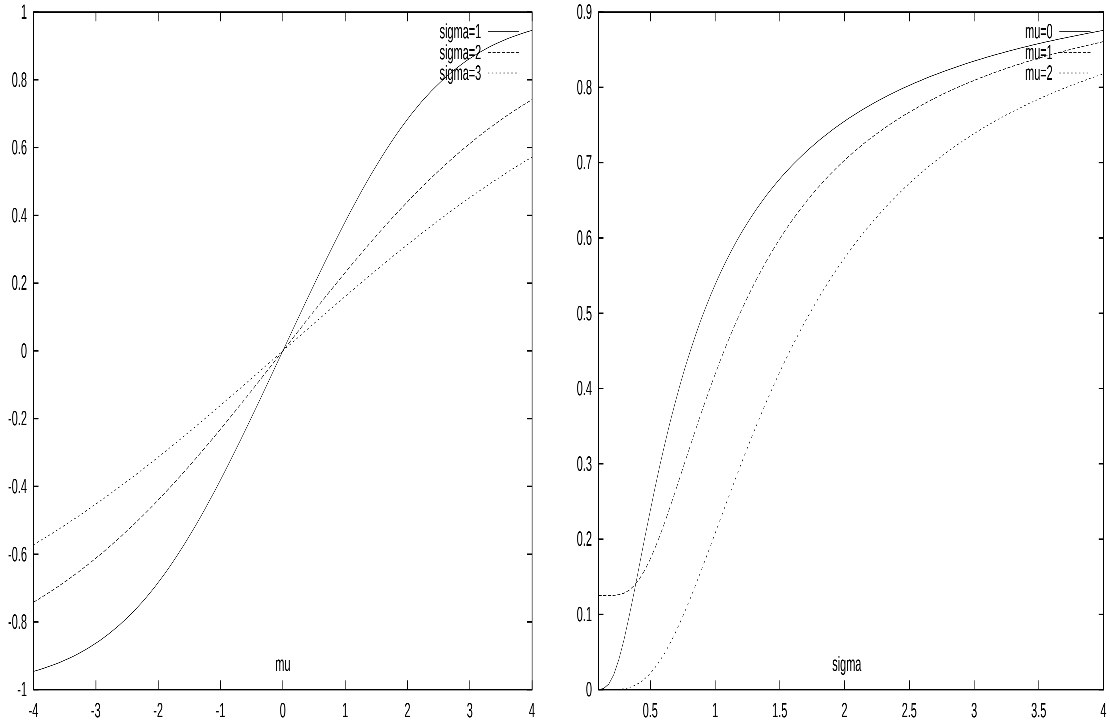

Figure 3: Dependence of $\mathbb{E}(\Delta p_i)$ on μ and σ^2 (left panel) and $\mathbb{V}(\Delta p_i)$ on μ and σ^2 (right panel) in the static model with trivariate distribution of Δp_i.

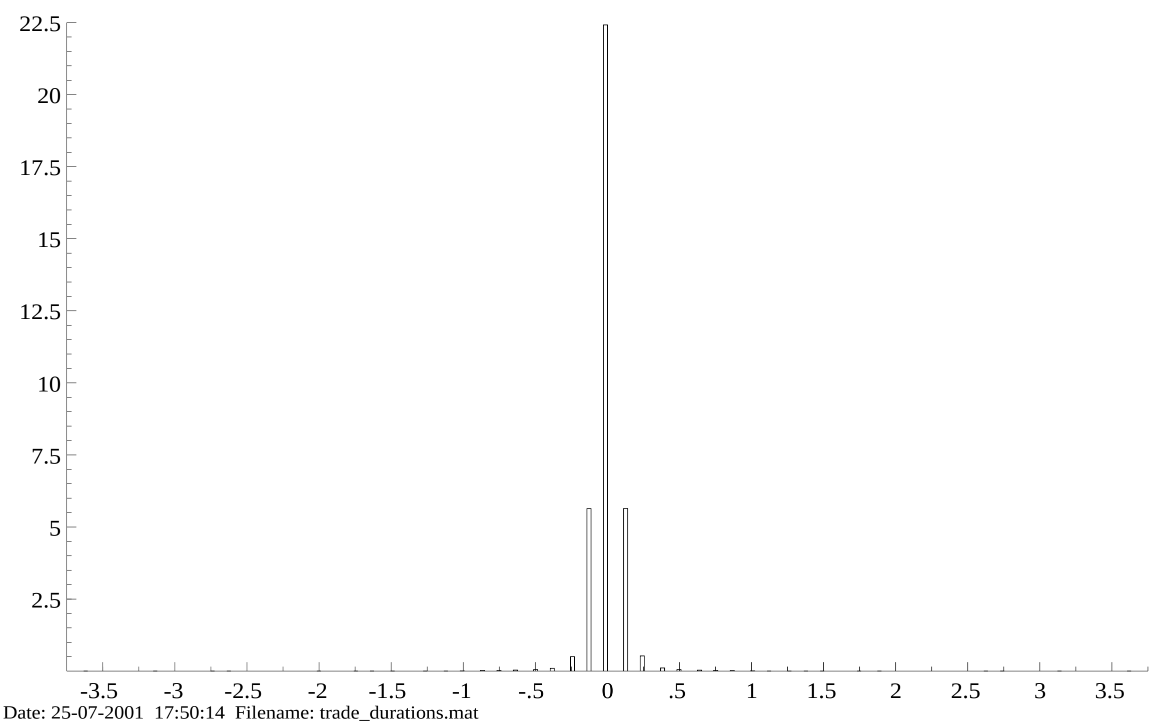

Figure 4: Unconditional distribution of Δp_i in IBM transaction data. Tick size equals to $\$\frac{1}{8}$.

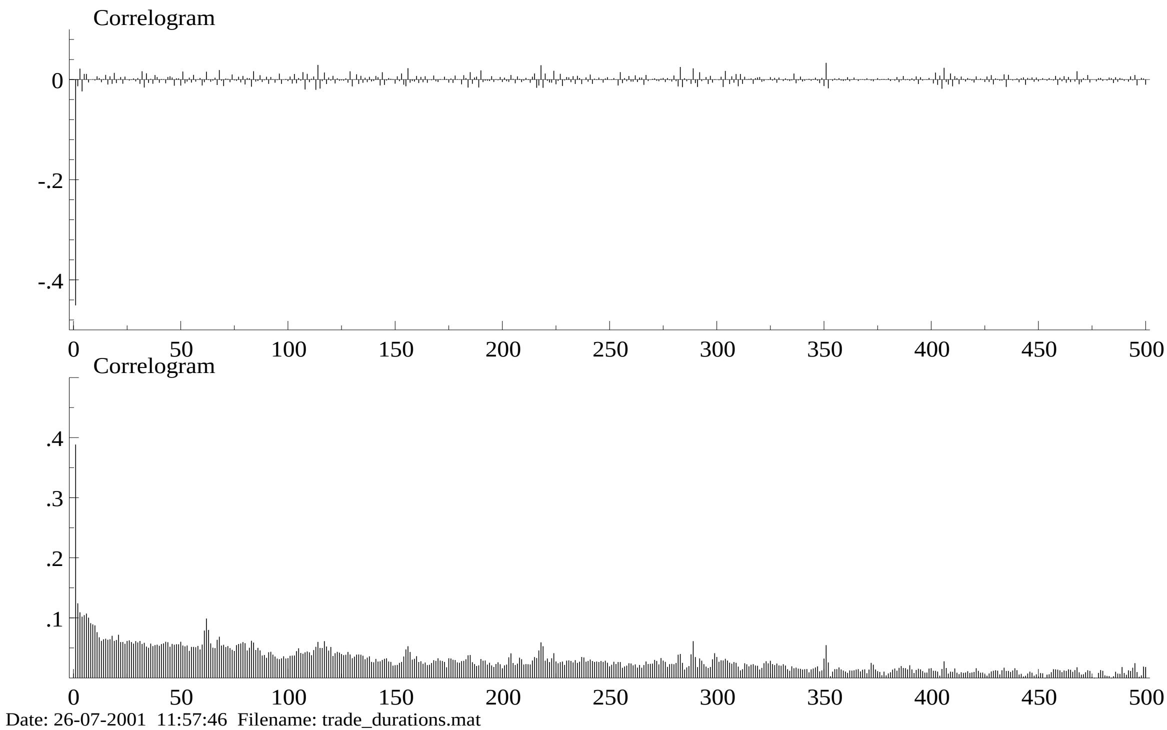

Figure 5: Correlograms of Δp_i (upper panel) and $|\Delta p_i|$ (lower panel) for IBM transaction data.

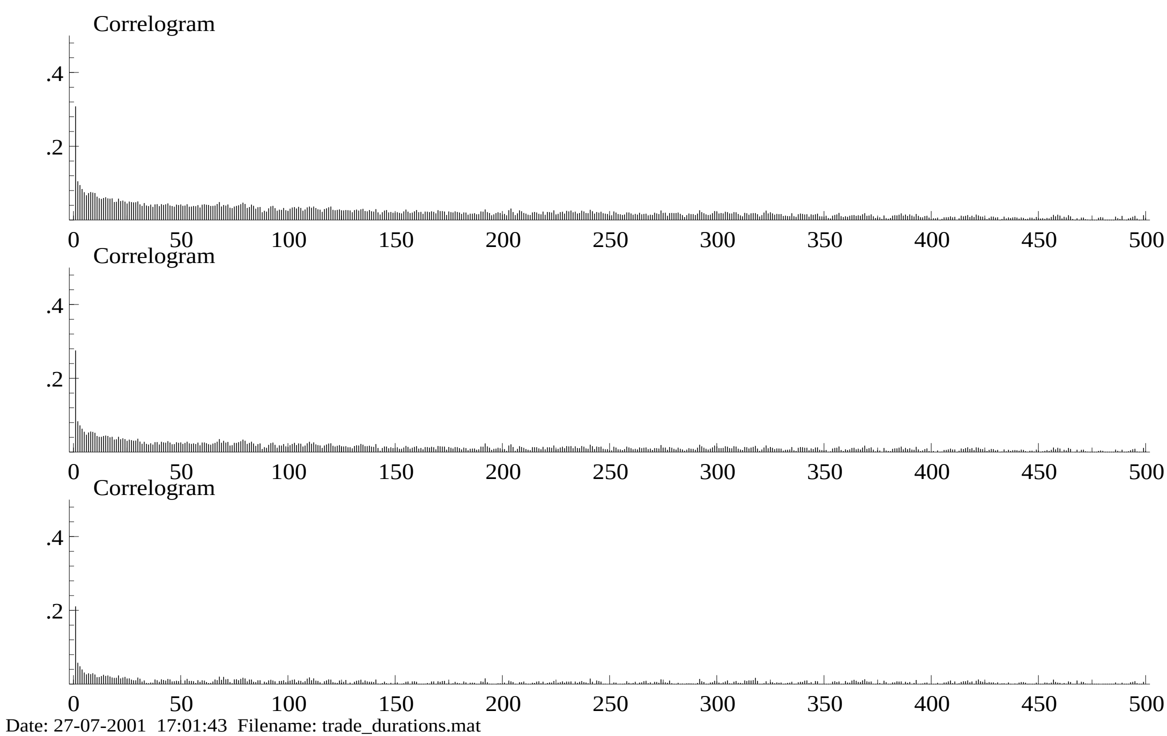

Figure 6: Correlograms of censored $|\Delta p_i|$ series: 7 states (upper panel), 5 states (middle panel) and 3 states (lower panel).

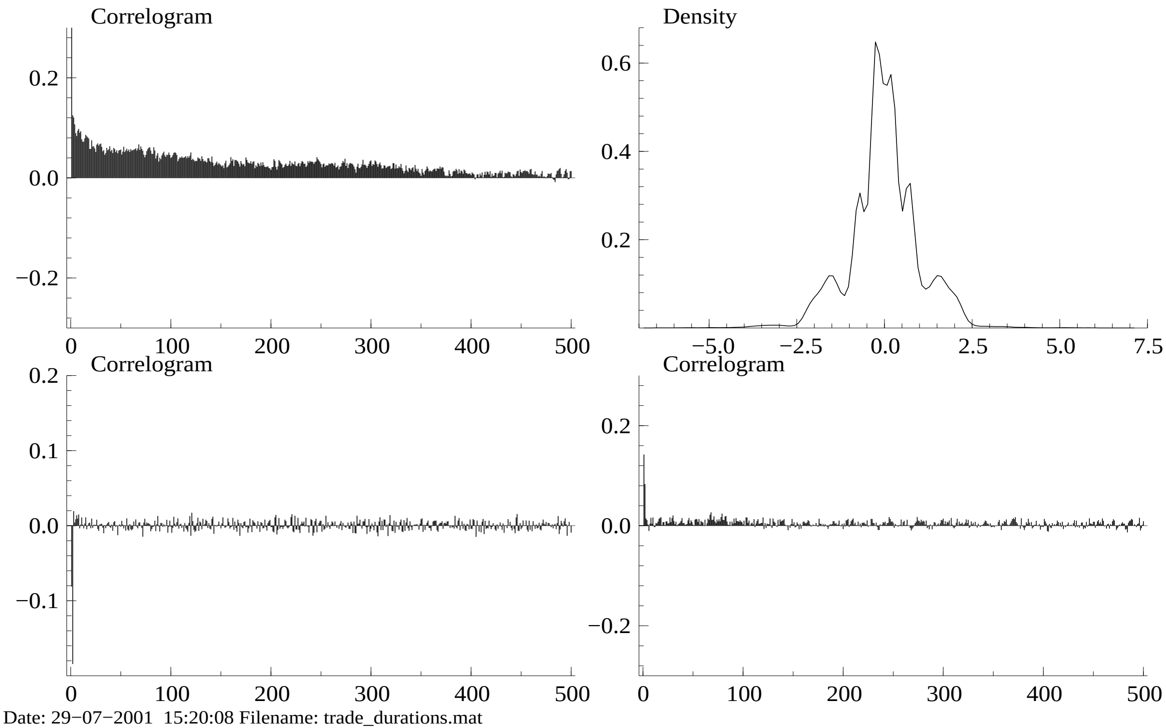

Figure 7: Residuals diagnostics for IV-GARCH model. Correlograms of $\hat{u}_i$ (lower left panel) and $|\hat{u}_i|$ (lower right panel). Estimated density of $\hat{u}_i$ (upper right panel) and correlogram of $|\Delta p_i|$ (upper left panel).

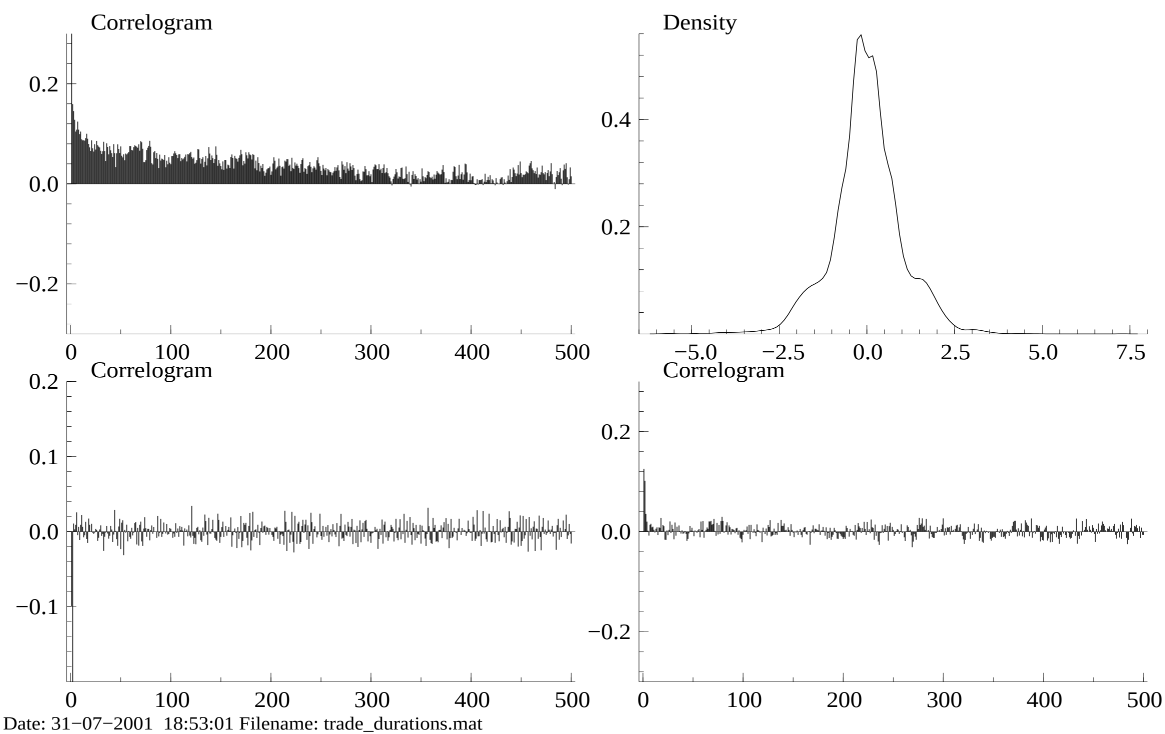

Figure 8: Residuals diagnostics for IV-FIARCH model. Correlograms of $\hat{u}_i$ (lower left panel) and $|\hat{u}_i|$ (lower right panel). Estimated density of $\hat{u}_i$ (upper right panel) and correlogram of $|\Delta p_i|$ (upper left panel).