GARCH models based on

Brownian Inverse Gaussian innovation processes

Gideon Griebenow BSc, MSc

Thesis submitted for the degree Doctor of Philosophy at
the Potchefstroom Campus of the North-West University

Promoter: Prof PJ de Jongh
Assistant Promoter: Prof JH Venter

September 2006

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES AND FIGURES

LIST OF TABLES AND FIGURES (CONTINUED)

LIST OF TABLES AND FIGURES (CONTINUED)

LIST OF TABLES AND FIGURES (CONTINUED)

LIST OF TABLES AND FIGURES (CONTINUED)

LIST OF TABLES AND FIGURES (CONTINUED)

"The teacher, if indeed wise, does not bid you to enter the house of their wisdom, but leads you to the threshold of your own mind." ~ Kahlil Gibran

PROLOGUE

Surviving a PhD is a lot like surviving life. You are born a very small person into a huge world of knowledge and you have to start learning very quickly. You need some milk to get you through the first few months until you can attempt to take your first uncertain bites of solid food. You have to crawl to where you want to be, since you will only much later learn to stand up and walk around. You will encounter many unfamiliar things – some of them inviting, some of them becoming mundane and some others just plain scary, but all of them essentially part of life. However, spend enough time with life and you may eventually learn to cut yourself a tiny piece of juicy steak, run a few meters at pace and familiarise yourself with some of the more worthwhile things it has to offer. And if you are lucky enough, you do not have to achieve all of this on your own.

I owe Professor Riaan de Jongh a great deal of gratitude for his support over the years that I have been privileged enough to know him. In addition to being both an amiable person and, as it were, a superior superior, he has always been willing to go out of his way in his professional and personal capacity whenever his assistance could make life easier for me. I am glad to have had him as a teacher in the lesson of life.

I am also deeply indebted to Professor Hennie Venter – without him this prologue may have been written either much earlier or never at all, both these scenarios being inferior to the eventual situation. I have learnt a great deal over the last few years and, although it was sometimes extremely difficult just to keep up with him, I could not have asked for a more competent teacher. Many seemingly perpetual problematic knots were disentangled by his ability and sheer persistence.

I would like to extend a special word of thanks to Professor Dries de Wet for the countless hours of study leave he granted me to conclude the final stubborn sections.

Finally, I am grateful for friends and family, opportunities and abilities, marvels and beauty. These are the things that have allowed me to come this far; to keep seeing the wonder of something I have seen a thousand times; to keep pondering the questions even if not for the sake of answers; to always keep taking joy from the journey we call life.

ABSTRACT

Title: GARCH models based on Brownian Inverse Gaussian innovation processes.

Key words: GARCH, Volatility, Brownian Inverse Gaussian Processes, Random Impact Factors, Time Dependence, Autocorrelation.

In classic GARCH models for financial returns the innovations are usually assumed to be normally distributed. However, it is generally accepted that a non-normal innovation distribution is needed in order to account for the heavier tails often encountered in financial returns. Since the structure of the normal inverse Gaussian (NIG) distribution makes it an attractive alternative innovation distribution for this purpose, we extend the normal GARCH model by assuming that the innovations are NIG-distributed. We use the normal variance mixture interpretation of the NIG distribution to show that a NIG innovation may be interpreted as a normal innovation coupled with a multiplicative random impact factor adjustment of the ordinary GARCH volatility. We relate this new volatility estimate to realised volatility and suggest that the random impact factors are due to a news noise process influencing the underlying returns process. This GARCH model with NIG-distributed innovations leads to more accurate parameter estimates than the normal GARCH model. In order to obtain even more accurate parameter estimates, and since we expect an information gain if we use more data, we further extend the model to cater for high, low and close data, as well as full intraday data, instead of only daily returns. This is achieved by introducing the Brownian inverse Gaussian (BIG) process, which follows naturally from the unit inverse Gaussian distribution and standard Brownian motion. Fitting these models to empirical data, we find that the accuracy of the model fit increases as we move from the models assuming normally distributed innovations and allowing for only daily data to those assuming underlying BIG processes and allowing for full intraday data.

However, we do encounter one problematic result, namely that there is empirical evidence of time dependence in the random impact factors. This means that the news noise processes, which we assumed to be independent over time, are indeed time dependent, as can actually be expected. In order to cater for this time dependence, we extend the model still further by allowing for autocorrelation in the random impact factors. The increased complexity that this extension introduces means that we can no longer rely on standard Maximum Likelihood methods, but have to turn to Simulated Maximum Likelihood methods, in conjunction with Efficient Importance Sampling and the Control Variate variance reduction technique, in order to obtain an approximation to the likelihood function and the parameter estimates. We find that this time dependent model assuming an underlying BIG process and catering for full intraday data fits generated data and empirical data very well, as long as enough intraday data is available.

OPSOMMING

Titel: GARCH modelle gebaseer op Brownse inverse Gaussiese prosesse.

Sleutelwoorde: GARCH, Volatiliteit, Brownse Inverse Gaussiese Prosesse, Stogastiese Impakfaktore, Tydafhanklikheid, Outokorrelasie.

In klassieke GARCH modelle vir finansiële opbrengste word daar gewoonlik aangeneem dat die innovasies normaal verdeel is. Swaarder sterte word egter dikwels in finansiële opbrengste waargeneem en 'n nie-normale innovasie-verdeling word benodig om voorsiening daarvoor te maak. Ons brei die normale GARCH model uit deur aan te neem die innovasies is normaal inverse Gaussies (NIG) verdeel, aangesien die struktuur van die NIG verdeling dit 'n aantreklike alternatiewe verdeling vir hierdie doel maak. Ons maak gebruik van die normale variansie mengsel interpretasie van die NIG verdeling om te wys dat 'n NIG innovasie geïnterpreteer kan word as 'n normale innovasie met 'n multiplikatiewe stogastiese impakfaktor aanpassing van die gewone GARCH volatiliteit. Hierdie stogastiese impakfaktore word gekoppel aan die nuusgebeure proses wat die onderliggende opbrengsproses beïnvloed. Meer akkurate parameterberamings word verkry deur middel van hierdie NIG-GARCH model. Om selfs groter akkuraatheid te verkry, en aangesien ons kan verwag dat meer data meer informasie sal bevat, brei ons die model verder uit om voorsiening te maak vir hoog, laag en sluitingsdata, asook volledige intradag data, in plaas van slegs daaglikse data. Dit word bewerkstellig deur die gebruik van die Brownse inverse Gaussiese (BIG) proses, wat 'n natuurlike uitbreiding van die inverse Gaussiese verdeling en standaard Brownse beweging is. Ons pas hierdie modelle op empiriese data en vind dat model akkuraatheid verhoog soos ons beweeg vanaf die modelle wat normaalverdeelde innovasies aanneem en slegs daaglikse data gebruik na die modelle wat onderliggende BIG prosesse aanneem en volledige intradag data gebruik.

Tog vind ons 'n problematiese resultaat, naamlik dat daar empiriese bewyse van tydafhanklikheid in die stogastiese impakfaktore voorkom. Dit beteken dat die nuusgebeure proses nie tydonafhanklik is soos wat ons aangeneem het nie. Inteendeel, ons kan verwag dat die stogastiese impakfaktore wel tydafhanklik is. Ten einde vir hierdie afhanklikheid voorsiening te maak brei ons die model nog verder uit. Die groter kompleksiteit wat met hierdie uitbreiding gepaard gaan noodsaak ons om staat te maak op Gesimuleerde Maksimum Aanneemlikhiedsmetodes (gekombineerd met variansie reduksie) in plaas van gewone Maksimum Aanneemlikheid, aangesien 'n eksplisiete uitdrukking vir die aanneemlikheidsfunksie nie beskikbaar is nie. Ons bevinding is dat hierdie tydafhanklike model, wat onderliggende BIG prosesse aanneem en voorsiening maak vir volledige intradag data, goeie passings op gegenereerde en empiriese data lewer, solank genoeg intradag data beskikbaar is.

CHAPTER 1

INTRODUCTION

1.1 Volatility

According to the *American Heritage Dictionary of the English Language*, the root of the word volatility is the Latin *volare*, which means "to fly". The *Oxford English Dictionary* defines the word volatile as "easily evaporated at normal temperature". This means that something which is volatile can suddenly and unexpectedly change, even in the absence of adverse conditions. In a financial context, investments that easily evaporate under normal conditions can lead to a loss of money and some very unhappy and unfortunate investors (with volatile moods).

The concept of volatility is probably the most researched phenomenon in financial markets. Asset price volatility is a measure of the uncertainty or randomness in the behaviour of the asset price, i.e. to what extent can we expect the asset price to move up or down randomly. Thus, volatility is a measure of the risk associated with the asset. In option pricing, the single-most important factor in determining the fair value of an option is the volatility in the price of the underlying asset. The opportunity for gaining or losing money by investing in such an option is rooted in the asset price volatility.

A measure of historical volatility can be obtained by observing the change in asset price over a specific time period. The classic historical volatility estimator is obtained by calculating the standard deviation of the asset prices over time around the mean price. Some alternative volatility estimators, which achieve higher efficiency (lower variance) by using high, low and open prices in addition to close prices, have been proposed by, for example, Parkinson (1980) and Rogers and Satchell (1991).

There is an abundance of evidence that shows that returns in the financial arena have non-constant volatility, which has led to the development of numerous non-constant volatility models. A jump-diffusion model (see Bates (1996) for an example) adds random jumps to the asset returns, while mixed distribution models (see Barndorff-Nielsen (1977) for an example) model the asset returns using a mixture of distributions. Stochastic volatility models (see Heston (1993) for an example) treat not only the returns, but also the volatility of the returns, as stochastic variables, which increases model complexity.

In the editorial of the special edition of the Journal of Applied Econometrics, Volume 17, dealing with volatility Franses and MacAleer (2002) motivates the influence and importance of volatility modelling as follows:

"Time-varying volatility is naturally relevant, topical, useful and interesting for modelling and forecasting financial data. The explosive growth of applications of econometrics to finance is due primarily to the increased availability of financial data, increased computer power and, of course, the greater interest in the performance of financial markets in current economic discussions. As important stock market fluctuations appear on the front pages of many leading newspapers, being economically and financially literate about trends and volatilities in stock prices or the uncertainty of trading on financial markets is mandatory in present-day business enterprises. Consequently, it is common to observe the substantial relevance of econometric methods in financial practices. Moreover, many financial firms and institutions rely on sophisticated econometric models to guide decision making, which demonstrates the relevance of financial econometrics for a multitude of practitioners in business and commerce."

We will restrict our attention to financial time series that describe the movement of financial series such as stock prices and financial indices. Stationarity is a desirable property in time series analysis - it is a minimal requirement for successful time series analysis, since the underlying process stays the same when the series is stationary, which means that well-developed analysis techniques can be employed (see Box and Jenkins (1976)).

However, price series are mostly non-stationary, as can be seen in Figure 1.1, which shows IBM closing prices for the 1000 trading days from 18 February 2000 to 12 February 2004. When comparing the period February 2000 to February 2002 with the period November 2002 to February 2004 it is clear that the IBM prices follow different trends for sustained periods over time. A standard technique used in time series analysis to transform a non-stationary series into one that is more stationary is to take first order differences. In financial time series analysis in particular it is common to work with the log prices as basic series and then take differences of these, getting what are commonly called the log-returns. These log-returns are then analysed instead of the prices themselves. We follow this practice in this thesis. Figure 1.1 also displays the log-returns for the IBM data. Clearly this series appears much more stationary over time.

One of the most important aspects to consider when defining a model for price level dynamics is how to describe price variability. The extent of the variation in the log-returns over time is the return volatility, which has been studied by many authors in the financial literature. Gigli (2002) gives a useful overview of the stylised facts regarding volatility that are regularly encountered in practice, including heavy tails, volatility clustering, mean reversion, leverage effects, information arrivals, implied volatilities and volatility smiles.

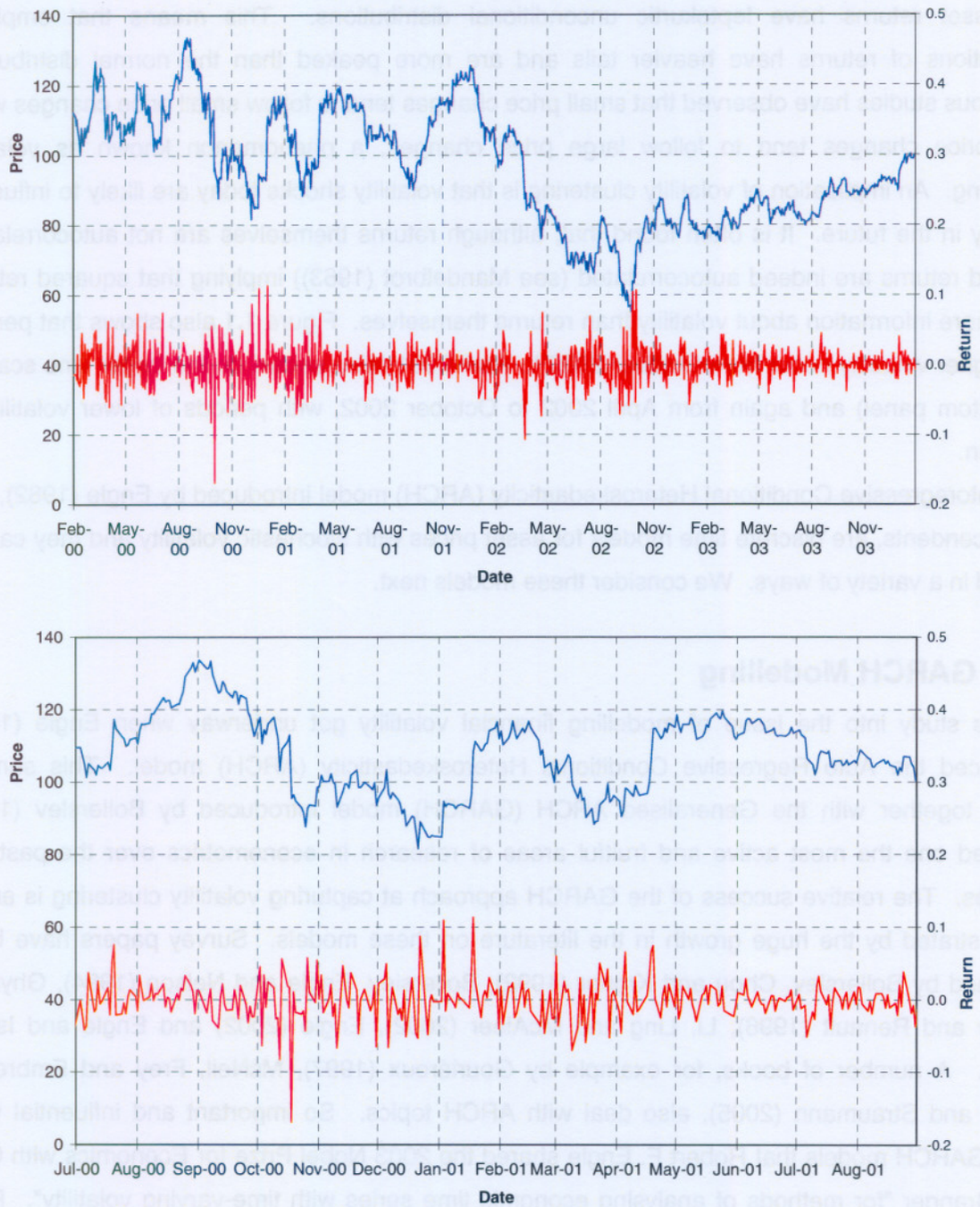

Figure 1.1: Volatility clustering in daily IBM returns.

In the context of this thesis the most relevant stylised facts that characterise financial time series are non-normality and volatility clustering of returns. Mandelbrot (1963) and Fama (1963) observe that asset returns have leptokurtic unconditional distributions. This means that empirical distributions of returns have heavier tails and are more peaked than the normal distribution. Numerous studies have observed that small price changes tend to follow small price changes while large price changes tend to follow large price changes, a phenomenon known as volatility clustering. An implication of volatility clustering is that volatility shocks today are likely to influence volatility in the future. It is often found that, although returns themselves are not autocorrelated, squared returns are indeed autocorrelated (see Mandelbrot (1963)) implying that squared returns carry more information about volatility than returns themselves. Figure 1.1 also shows that periods with higher volatility occur from October 2000 to April 2001 (also shown on the larger time scale in the bottom panel) and again from April 2002 to October 2002, with periods of lower volatility in between.

The Autoregressive Conditional Heteroskedasticity (ARCH) model introduced by Engle (1982), and its descendents, are discrete time models for asset prices with stochastic volatility and they can be defined in a variety of ways. We consider these models next.

1.2 GARCH Modelling

Serious study into the issue of modelling financial volatility got underway when Engle (1982) introduced the Auto Regressive Conditional Heteroskedasticity (ARCH) model. This seminal paper, together with the Generalised ARCH (GARCH) model introduced by Bollerslev (1986) triggered one the most active and fruitful areas of research in econometrics over the past two decades. The relative success of the GARCH approach at capturing volatility clustering is amply demonstrated by the huge growth in the literature on these models. Survey papers have been published by Bollerslev, Chou and Kroner (1992), Bollerslev, Engle and Nelson (1994), Ghysels, Harvey and Renault (1996), Li, Ling and McAleer (2002), Engle (2002) and Engle and Ishida (2002). A number of books, for example by Gouriéroux (1997), McNeil, Frey and Embrechts (2005) and Straumann (2005), also deal with ARCH topics. So important and influential were these GARCH models that Robert F. Engle shared the 2003 Nobel Prize for Economics with Clive W.J. Granger "for methods of analysing economic time series with time-varying volatility". From the surveys mentioned above, it is clear that numerous issues are considered in the vast amount of existing GARCH literature. As these surveys are quite recent, we will not attempt to review the issues addressed in the existing literature. Instead, we will focus mainly on matters relating to specification of the innovation distribution in GARCH models.

Some notation will be helpful for this purpose. If we denote the daily closing prices of a financial instrument by $p_1, p_2, \ldots$ we can write the daily log return on day n as $Y_n = \log(p_n / p_{n-1})$ and describe the returns series using a basic GARCH model of the form

$$Y_n = \mu_n + \sqrt{h_n} X_n \text{ for } n = 1, 2, \ldots . \tag{1.2.1}$$

Here μ_n represents an expected (or structural) component, h_n is the volatility and X_n is the innovation at time n, assumed to have zero expectation and unit variance. Both μ_n and h_n are assumed to depend at most on the past observations $Y_1, \ldots, Y_{n-1}$ and the X_n's are assumed independent and identically distributed (iid). In Section 3.2 we give a more detailed discussion and some extensions, as well as the specifications of μ_n and h_n.

GARCH models are traditionally fitted to data by means of quasi-maximum likelihood (QML), i.e. the innovation X_n is assumed to have a standard normal distribution and maximum likelihood estimation (MLE) is carried out. However, the inability of GARCH models with the assumption of normally distributed innovations to fully account for the mass in the tails of empirically estimated daily innovations, is also generally well recognised (see Forsberg and Bollerslev (2002)). Several alternative innovation distributions have been proposed, for example the t-distribution by Bollerslev (1987) and McNeil and Frey (2000), the general error distribution by Nelson (1991) and, a recent strong favourite, the normal inverse Gaussian (NIG) distribution by Barndorff-Nielsen (1997).

A line of approach towards dealing with the non-normality issue dates back to the mixture-of-distributions-hypothesis (MDH) of Clark (1973) according to which a suitable variance mixture of normal distributions is required for the innovation distribution. This approach was recently given new impetus by the work of Forsberg and Bollerslev (2002), who used the notion of realised volatility (see Andersen et al. (2000-2001) or Section 3.1.3 below) to guide in the selection of a mixing distribution. They find that the distribution of realised volatility standardised by a GARCH type volatility estimate is well approximated by an inverse Gaussian (IG) distribution. This ties in with results that will be discussed in the next section and suggests that the NIG distribution should be a particularly appropriate choice for the innovation distribution. This elucidates the findings of authors such as Venter and de Jongh (2004) that using the NIG distribution in this context often leads to more efficient estimation of model parameters and related quantities such as volatility and risk measures. This thesis develops the consequences of using the NIG distribution for the innovations. The next section provides an introductory overview of the approach and contributions reported here.

1.3 Outline of the Thesis

If the innovation X_n in (1.2.1) is assumed to be NIG-distributed, then it may be written in the form $X_n = \sqrt{W_n} Z_n$ where Z_n is normally distributed and W_n is IG-distributed. A more precise formulation and details on this assertion are provided in Chapter 2. Accepting it for now, equation (1.2.1) may be written as

$$Y_n = \mu_n + \sqrt{h_n}\sqrt{W_n} Z_n = \mu_n + \sqrt{h_n W_n} Z_n \,. \qquad (1.3.1)$$

The second expression here shows that the model may be thought of as having a normal innovation term Z_n and a modified volatility $h_n W_n$. This interpretation is intuitively meaningful. Since the GARCH volatility h_n can only depend on past returns, it cannot incorporate newly arriving information that may shock today's volatility to a level different from the value h_n predicted from the past. Taking W_n to represent the effect of such a shock, $h_n W_n$ would represent the more relevant actual volatility experienced today. Thus, the NIG innovation specification allows for meaningful interplay between innovation and volatility. This is the approach that will be explored in great detail in this thesis.

Up to this point we only dealt with what may be called the "daily (D) data case", i.e. we have at most daily open and close prices and thus daily returns available and daily volatility is studied in this context. However, intraday events also have an important bearing on daily volatility. For example, suppose the price opens the day at \$100, then plummets to \$90 at noon and recovers to close at \$99 at the end of the day. To the casual observer reading successive closing prices in the newspaper it appears to have been a normal trading day, but an active trader actually experienced great volatility. However, this intraday price volatility would bear no consequence on a traditional daily GARCH model and this seems unrealistic. We would expect to get some improvement if we used the high and low prices in addition to the open and close prices and even more improvement if we could involve the complete set of intraday prices in the model. Clearly it is important to integrate as much intraday data as are available into volatility models in order to obtain more realistic models.

Indeed, Engle (2002) identifies high-frequency volatility models as a new frontier for ARCH models and "… [an area] where substantial research can be expected over the next few years." He goes on to say that "[t]he study of volatility models within the day is in its infancy yet is a natural extension of the daily models examined so widely."

The availability and reliability of financial data has come a long way over the last few decades. In years past financial data consisted mainly of daily observations (for example daily close prices)

occasionally accompanied by extreme observations (high and low prices). Naturally, the majority of models developed in these times allowed for only the available daily data. The widespread increase in availability and dependability of higher-frequency data of the past few years has had a significant impact on research in the financial arena, especially the modelling of daily and intraday volatility.

To provide an overview of our contributions in this regard, we extend the notation above to the intraday situation. Denote the price at time t on day n by $p_{t,n}$, where we measure time in terms of a fraction of the trading day so that $0 \le t \le 1$, $t=0$ corresponds to opening and $t=1$ to closing time. The return up to time t on day n can then be denoted by $Y_{t,n} = \log\left(p_{t,n}/p_{0,n}\right)$. Consider the case where we observe the open, high, low and close prices (i.e. $p_{0,n}$, $\max_t p_{t,n}$, $\min_t p_{t,n}$ and $p_{1,n} = p_{0,n+1}$). We call this the "high, low, close (HLC) case". Associated with these prices are the total return for the day $Y_n = Y_{1,n}$, the high return defined as $A_n = \log\left(\max_t p_{t,n}/p_{0,n}\right) = \max_t Y_{t,n}$ and the low return defined as $B_n = \log\left(p_{0,n}/\min_t p_{t,n}\right) = -\min_t Y_{t,n}$. In principle, once a model is specified for the intraday return stochastic processes $Y_{t,n}, 0 \le t \le 1$, it is possible to compute the joint distributions of (Y_n, A_n, B_n) and hence to use high, low and close prices to fit the model via likelihood methods. A model suggested by (1.3.1) is to assume that

$$Y_{t,n} = \mu_n t + \sqrt{h_n W_n} Z_{t,n} \qquad (1.3.2)$$

with $Z_{t,n}$ standard Brownian motion in t, independent over n and μ_n, h_n and W_n as in (1.3.1). This model nests (1.3.1) but allows us to incorporate additional information coming from the high, low and close prices in the fitting process.

Consider next the case where we break up the trading day into I time intervals of equal length $\Delta = 1/I$ each and we observe the prices $p_{i\Delta,n}$ at the times $i\Delta, i = 0,1,\ldots,I$. We call this the "intraday (ID) case". Associated with these prices are the returns $Y_{i\Delta,n}, i = 0,1,\ldots,I$. Again, in principle, once a model is specified for $Y_{t,n}, 0 \le t \le 1$, it is possible to compute the joint distributions of $\{Y_{i\Delta,n}, i = 0,1,\ldots,I\}$ and hence to fit the model via likelihood methods using this intraday price data. Again, the model (1.3.2) can be used in this context, hence allowing us to incorporate full intraday data if it is available.

In both cases the structural part μ_n and the GARCH volatility part h_n remain as in the daily GARCH models, but since more informative data is used in the fitting process more accurate estimation should be possible. This generally turns out to be the case. In the intraday case it turns out that $Y_{1,n} \equiv Y_n$ and $R_n = \sum_{i=1}^{I} \left(Y_{i\Delta,n} - Y_{(i-1)\Delta,n} \right)^2$ constitute sufficient statistics on which to base model fitting. The statistic R_n has become important recently as a possible measure of daily volatility based on intraday data and is called the "realised volatility". It was introduced by Taylor and Xu (1997) and Andersen and Bollerslev (1998) and was studied extensively in Andersen et al. (2003). The work reported in this thesis clarifies the relation between daily GARCH volatility estimates and realised volatility. This may be summarised in the approximate relation $R_n \approx h_n W_n$ which states that realised volatility follows GARCH volatility after adjusting the latter by the daily random impact factor W_n. These W_n's are latent (unobserved) variables but they can be estimated in the intraday case, which makes it practical to verify the relation between realised and GARCH volatilities. We report empirical evidence supporting this relation. The finding of Forsberg and Bollerslev (2002) referred to in the previous section can be understood from this relation. Since $R_n \approx h_n W_n$ we have $R_n / h_n \approx W_n$ and since W_n is assumed to be inverse Gaussian (IG) distributed, this is the finding that realised volatility standardised by GARCH volatility is IG-distributed.

GARCH models of the form (1.2.1) always assume that the innovations X_n are independent over n. When we represent X_n in the form $\sqrt{W_n} Z_n$ in the case of (1.3.1), or more generally when we work with innovations processes $X_{t,n} = \sqrt{W_n} Z_{t,n}$ in (1.3.2), we assumed that both Z_n (or $Z_{t,n}$) and W_n are independent over n. However, when testing the independence assumption for the W_n's in the case of intraday empirical data, it turned out that this was not a realistic assumption. This strengthens our interpretation that the W_n's represent impact factors emanating from news events: major news events are likely to have longer lasting effects on volatility than just one day. Unfortunately, such dependence has serious technical consequences for model fitting and analysis using likelihood methods: computing the likelihood function under dependence is much more complex! We addressed this problem by means of efficient importance sampling (EIS) and it turns out that parameter estimates found in this way may differ from those based on the simplifying (but erroneous) independence assumption.

More detailed formulations and discussions of the models, data cases and results obtained above are given in Chapters 2 to 4.

1.4 Structure of the Thesis

Chapter 2 contains the distribution and stochastic process theory needed in Chapter 3, where it is applied to the various GARCH models developed in this thesis. On first reading Chapter 2 the topics touched on may appear somewhat lacking in motivation. We request that the reader be patient in this regard: motivation will become clear when the results are applied in Chapter 3. An alternative exposition could have been to develop the necessary distribution theory as and when required while working with the GARCH models. However, we thought it preferable to do the distribution theory in the simpler and notationally cleaner context where the GARCH details were not present.

Sections 2.1 and 2.2 review the generalised inverse Gaussian (GIG), inverse Gaussian (IG) and unit inverse Gaussian (UIG) distributions; the UIG distribution is especially important in this thesis and plays a major role as the distributional model for the random impact factors W_n in the rest of the thesis. Section 2.3 introduces a standardised form of the normal inverse Gaussian (NIG) distribution, referred to as the standard normal inverse Gaussian (SNIG) distribution, which is a mean-variance mixture of the normal distribution with the UIG distribution as mixing distribution. This serves as the innovation distribution in the case where only daily close prices are available. Section 2.4 and 2.5 introduce the standard chi-squared inverse Gaussian (SCIG) and bivariate standard normal chi-squared inverse Gaussian (SNCIG) distributions, which are needed for likelihood inference in the intraday data case. Section 2.6 introduces the t inverse Gaussian (TIG) distribution, which can be used in residual analysis contexts. Section 2.7 defines the Brownian inverse Gaussian (BIG) process, which serves as the innovation process in the models for the high, low, close data and intraday data cases. We derive expressions for the joint densities of the various functionals of the BIG process required for likelihood inference in these cases. Section 2.8 deals with the topic of generating independent random variables from IG related distributions. This is needed in order to be able to generate data sets with which to test our programs, as well as for performing residual analysis.

After the Chapter 3 outline given in Section 3.1, we present several GARCH-type models for modelling financial time series in Section 3.2. The main differences between the models are the frequency of data that each model caters for and the innovation distribution assumed for each model, i.e. each model caters for either daily close prices, daily high, low and close prices, or intraday prices, and assumes either normally distributed innovations (or underlying BM processes) or NIG-distributed innovations (or underlying BIG processes). In Section 3.3 we present the results of fitting these models to several data sets. We use generated data sets, to test our programs and illustrate the underlying theory, as well as empirical data sets. We find that models catering for

intraday data, and assuming underlying BIG processes, provide much more accurate parameter estimates than models catering for only daily data, and assuming underlying BM processes. However, we also find empirical evidence of time dependence in some of the model variables (specifically the random impact factors) which we assumed to be independent over time. Therefore, in Section 4.2, we present a time-dependent version of the intraday GARCH model with an underlying BIG process. Section 4.3 presents results of fitting this time-dependent model to both generated and empirical data sets. We find that this time-dependent model fits empirical data well as long as enough intraday data is available. We also compare these results with the results obtained in Chapter 3 in order to assess the reliability of the results obtained from fitting the models assuming independence to data where dependence is actually present. Chapter 5 summarises our findings and presents topics for possible future research.

In Section A.1 of the Appendix we briefly review Quantile-Quantile (Q-Q) and Probability Integral Transform (PIT) plots, which we use throughout Sections 3.3 and 4.3 to assess the goodness of fit of various models. A disadvantage of the Q-Q and PIT plot based assessment of the quality of fit is that it tends to be rather informal in the sense that no numerical measure of fit is computed to which a p-value can be attached. However, the use of these plots is widespread in the relevant literature (see e.g. Forsberg (2002)) and we shall follow this tradition. Section A.2 is a continuation of Section 4.2 and deals with approximating the autocorrelation in sequences of dependent random variables with UIG marginal distributions. It also proposes an algorithm for generating such sequences, which can be used for validation purposes. Finally, in Section A.3 we discuss some of the issues concerning the computer programs we wrote and supply some of the source code. The methodology we followed throughout regarding the practical fitting of models will become clear when examining the source code and accompanying comments. For all the programs, and partial programs, we used generated data sets to test our programs and derivations before running the programs with empirical data as input.

1.5 Terminology and Notation Conventions

The following abbreviations and acronyms are used throughout the thesis:

- AR: Autoregresive.
- ARCH: Autoregressive conditional heteroskedasticity.
- ARMA: Autoregressive moving average.
- BM: Brownian motion.
- BIG: Brownian inverse Gaussian.
- DW: Durbin-Watson.

- EIS: Efficient importance sampling.
- FORTRAN®: A high-level mathematical programming language (stands for Formula Translation).
- GARCH: Generalised autoregressive conditional heteroskedasticity.
- GIG: Generalised inverse Guassian.
- IG: Inverse Gaussian.
- iid: Independently and identically distributed.
- JSE: Johannesburg stock exchange.
- LM: Lagrange multiplier.
- LN: Log-normal.
- MDH: Mixture-of-Distributions-Hypothesis.
- ML: Maximum likelihood.
- MLE: Maximum likelihood estimation.
- MSE: Mean squared error.
- NAGARCH: Non-linear asymmetric generalised autoregressive conditional heteroskedasticity.
- NIG: Normal inverse Gaussian.
- NMC: Natural Monte Carlo.
- PIT: Probability integral transform.
- PML: Pseudo maximum likelihood (see QML).
- Q-Q: Quantile-Quantile.
- QML: Quasi maximum likelihood (see PML).
- RV: Realised volatility.
- SARV: Stochastic autoregressive volatility.
- SAS®: A statistical and matrix handling programming language (stands for Statistical Analysis System).
- SCIG: Standard chi-squared inverse Gaussian.
- SML: Simulated maximum likelihood.
- SNCIG: Standard normal chi-squared inverse Gaussian.
- SNIG: Standard normal inverse Gaussian.
- SSD: Sum of squared deviations.
- TIG: t inverse Guassian.
- UIG: Unit inverse Guassian.
- ULN: Unit log-normal.

We will adhere to the following notation conventions:

- Expressions are numbered according to the (second-level) section in which they appear. For instance, the fifth expression in Section 2.3, which could actually appear in one of the third-level sections (e.g. Section 2.3.2), would be numbered (2.3.5).
- K_λ is the modified Bessel function of third order and index λ (see Section 1.6).
- We will use $\log(x)$ to indicate the natural logarithm of x.

Finally, we list a number of formulae and notation conventions relating to densities of random variables for ease of reference:

- Whenever we have a continuously distributed random variable X we shall indicate its probability density function in the argument x by $f_X(x)$ and its distribution function in the argument x by $F_X(x)$. Alternatively, if the distribution of X is indicated by, for example, $X \sim N(0,1)$, we shall indicate the density function in the argument x by $f_{N(0,1)}(x)$ and the distribution function in the argument x by $F_{N(0,1)}(x)$.
- The normal distribution with mean μ and variance σ^2 has density

$$f_{N(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right). \tag{1.5.1}$$

- We shall indicate the standard normal density function in the argument z by

$$\varphi(z) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right) = f_{N(0,1)}(z) \tag{1.5.2}$$

and the standard normal distribution function in the argument z by

$$\Phi(z) = \int_{-\infty}^{z} f_{N(0,1)}(x)\,dx = F_{N(0,1)}(z) \tag{1.5.3}$$

- If $Y = aX + b$, with $a \neq 0$ and b real numbers, it follows that

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right). \tag{1.5.4}$$

- If $Y_1 = aX_1 + b$ and $Y_2 = cX_2 + d$ with $a \neq 0$, $c \neq 0$, b and d real numbers, it follows that

$$f_{Y_1,Y_2}(y_1,y_2) = \frac{1}{|ac|} f_{X_1,X_2}\left(\frac{y_1-b}{a},\frac{y_2-d}{c}\right). \tag{1.5.5}$$

1.6 Bessel functions

Many of the distributions dealt with in this thesis are defined in terms of Bessel functions of the third order $K_\lambda(x)$ for integral and half-integral values of λ. Here we list formulae required to compute these functions. These formulae may be found in e.g. Abramowitz and Stegun (1965). Since

$$K_{-\lambda}(x) = K_\lambda(x) \tag{1.6.1}$$

we need only consider $\lambda \geq 0$. For $\lambda = n + 1/2$ with n an integer, we may use the relations

$$K_{n+\frac{1}{2}}(x) = K_{\frac{1}{2}}(x)\left[1 + \sum_{i=1}^{n} \frac{(n+i)!}{(n-i)!\,i!} \frac{1}{(2x)^i}\right] \tag{1.6.2}$$

$$\text{while } K_{\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2}} x^{-\frac{1}{2}} \exp(-x)\ . \tag{1.6.3}$$

We also use the recursive relation

$$K_\lambda(x) = \frac{2(\lambda-1)}{x} K_{\lambda-1}(x) + K_{\lambda-2}(x) \tag{1.6.4}$$

to express any K_λ (for λ integer of half-integer values) in terms of K_1, $K_{1/2}$ and K_0. $K_{1/2}$ is explicitly given in (1.6.3) but K_1 and K_0 must be calculated using approximations such as those given by formulae (9.8.5) and (9.8.7) of Abramowitz and Stegun (1965). Many software packages include suitable subroutines for obtaining these approximations, which ease the use of these functions in numerical work. In FORTRAN these subroutines are given by DBSK1 and DBSK0.

CHAPTER 2

DISTRIBUTIONS AND STOCHASTIC PROCESSES RELATED TO INVERSE GAUSSIAN DISTRIBUTIONS

In this chapter we focus on some distribution and stochastic process theory in preparation for implementations in financial time series analysis in Chapters 3 and 4. We present properties of distributions related to the inverse Gaussian (IG) distribution, as well as stochastic processes based on Brownian motion (BM) and the IG distribution.

2.1 Definition of the GIG Distribution

The generalised inverse Gaussian (GIG) distribution was proposed more than 50 years ago by Good (1953). It has been used since then for many purposes. Among others, Sichel (1974, 1975) used it to construct mixtures of Poisson distributions, and Barndorff-Nielsen (1977, 1978) used it to obtain the generalized hyperbolic distribution as a mixture of normal distributions. Blæsild (1978) computed moments and cumulants and also considered the shape of the density. Jørgensen (1982) presents the statistical properties of the GIG distribution, working with its density in the form

$$\frac{(\psi/\chi)^{\frac{\lambda}{2}}}{2K_\lambda\left(\sqrt{\chi\psi}\right)} x^{\lambda-1}\exp\left(-\frac{1}{2}\left(\frac{\chi}{x}+\psi x\right)\right) \text{ for } x>0,\ \lambda\in\mathbb{R},\ (\chi,\psi)\in\Theta_\lambda$$

$$\text{where } \Theta_\lambda=\begin{cases}\{(\chi,\psi):\chi\geq 0,\psi>0\} \text{ if } \lambda>0\\ \{(\chi,\psi):\chi>0,\psi>0\} \text{ if } \lambda=0\\ \{(\chi,\psi):\chi>0,\psi\geq 0\} \text{ if } \lambda<0\ .\end{cases} \tag{2.1.1}$$

Some special cases of (2.1.1) are the gamma distribution ($\chi=0,\lambda>0$), the reciprocal gamma distribution ($\psi=0,\lambda<0$) and the hyperbola distribution ($\lambda=0$). Jørgensen (1982) also introduces the parameters $\omega=\sqrt{\chi\psi}$ as a location parameter (for fixed λ) and $\eta=\sqrt{\chi/\psi}$ as a scale parameter. Following, among others, Rydberg (1997) and Barndorff-Nielsen et al. (1992), we choose to work with an alternative parameterisation. We keep λ, but let $\delta=\sqrt{\chi}$ and $\gamma=\sqrt{\psi}$. With these parameters the $GIG(\lambda,\delta,\gamma)$ distribution has density

$$f_{GIG(\lambda,\delta,\gamma)}(w)=\left(\frac{\gamma}{\delta}\right)^{\lambda}\frac{w^{\lambda-1}}{2K_{\lambda}(\delta\gamma)}\exp\left(-\frac{1}{2}\left(\delta^2 w^{-1}+\gamma^2 w\right)\right) \text{ for } w>0,\ (\delta,\gamma)\in\Theta_{\lambda}$$

$$\text{where } \Theta_{\lambda}=\begin{cases}\{(\delta,\gamma):\delta\geq 0,\gamma>0\} \text{ if } \lambda>0\\ \{(\delta,\gamma):\delta>0,\gamma>0\} \text{ if } \lambda=0\\ \{(\delta,\gamma):\delta>0,\gamma\geq 0\} \text{ if } \lambda<0\ .\end{cases} \tag{2.1.2}$$

Comparing δ and γ with the parameters ω and η, we see that $\delta\gamma$ is a location parameter (for fixed λ) and δ/γ is a scale parameter. Since (2.1.2) integrates to 1 we get the following integral identity

$$\int_0^{\infty} w^{\lambda-1}\exp\left(-\frac{1}{2}\left(\delta^2 w^{-1}+\gamma^2 w\right)\right)dw=2\left(\frac{\delta}{\gamma}\right)^{\lambda}K_{\lambda}(\delta\gamma)\ . \tag{2.1.3}$$

Using (2.1.2) and (2.1.3) we can write

$$\begin{aligned}E\left(W^k\right)&=\left(\frac{\gamma}{\delta}\right)^{\lambda}\frac{1}{2K_{\lambda}(\delta\gamma)}\int_0^{\infty}w^{k+\lambda-1}\exp\left(-\frac{1}{2}\left(\delta^2 w^{-1}+\gamma^2 w\right)\right)dw\\ &=\frac{2\left(\frac{\gamma}{\delta}\right)^{\lambda}\left(\frac{\delta}{\gamma}\right)^{k+\lambda}K_{k+\lambda}(\delta\gamma)}{2K_{\lambda}(\delta\gamma)}\\ &=\delta^k K_{k+\lambda}(\delta\gamma)/\gamma^k K_{\lambda}(\delta\gamma)\ .\end{aligned} \tag{2.1.4}$$

Next, we consider some special cases of the GIG distribution.

2.2 The IG and UIG Distributions

2.2.1 Definition

When $\lambda=-1/2$ in the GIG distribution presented above we get the inverse Gaussian distribution, denoted by $IG(\delta,\gamma)$. In view of (1.6.1) and (2.1.2) its density is

$$f_{IG(\delta,\gamma)}(w)=\left(\frac{\gamma}{\delta}\right)^{-1/2}\frac{w^{-3/2}}{2K_{1/2}(\delta\gamma)}\exp\left(-\frac{1}{2}\left(\delta^2 w^{-1}+\gamma^2 w\right)\right) \text{ for } w>0. \tag{2.2.1}$$

Chhikara and Folks (1989) present the theory, methodology and application of the $IG(\delta,\gamma)$ distribution. We are interested in a special case of the IG distribution. Taking $\delta=\gamma=\psi>0$ leads to what we call the "unit inverse Gaussian" distribution, denoted by $UIG(\psi)$, with density (using (1.6.3))

$$f_{UIG(\psi)}(w) = \frac{w^{-3/2}}{2K_{1/2}(\psi^2)} \exp\left(-\frac{1}{2}\psi^2\left(w^{-1}+w\right)\right)$$
$$= \frac{\psi e^{\psi^2}}{\sqrt{2\pi}} w^{-3/2} \exp\left(-\frac{1}{2}\psi^2\left(w^{-1}+w\right)\right) \text{ for } w>0\,. \tag{2.2.2}$$

The $UIG(\psi)$ distribution function can be derived as a special case of the IG distribution function given by Chhikara and Folks (1974), and can be written as

$$F_{UIG(\psi)}(x) = \Phi\left(\psi\left(\sqrt{x}-\frac{1}{\sqrt{x}}\right)\right) + \exp\left(2\psi^2\right)\Phi\left(\psi\left(-\sqrt{x}-\frac{1}{\sqrt{x}}\right)\right)\,. \tag{2.2.3}$$

2.2.2 Properties

If W is $UIG(\psi)$-distributed then it follows from (2.1.4) that $E\left(W^k\right) = K_{k-1/2}\left(\psi^2\right)\Big/K_{1/2}\left(\psi^2\right)$ and in particular, using (1.6.2) and (1.6.3),

$$\begin{aligned} E(W) &= 1 \\ E\left(W^2\right) &= 1+\frac{1}{\psi^2} \\ E\left(W^3\right) &= 1+\frac{3}{\psi^2}+\frac{3}{\psi^4} \\ E\left(W^4\right) &= 1+\frac{6}{\psi^2}+\frac{15}{\psi^4}+\frac{15}{\psi^6} \end{aligned} \tag{2.2.4}$$

so that

$$\begin{aligned} Var(W) &= E\left(W^2\right)-\left(E(W)\right)^2 \\ &= 1/\psi^2 \\ Skew(W) &= \frac{E\left(W^3\right)-3E(W)E\left(W^2\right)+2\left(E(W)\right)^3}{Var(W)^{3/2}} \\ &= \psi^3+3\psi^{-1} \\ Kurt(W) &= \frac{E\left(W^4\right)-4E(W)E\left(W^3\right)+6\left(E(W)\right)^2E\left(W^2\right)-3\left(E(W)\right)^4}{Var(W)^2}-3 \\ &= 15\psi^{-2}\,. \end{aligned} \tag{2.2.5}$$

Figure 2.1 shows several $UIG(\psi)$ densities for different values of ψ. We can see that larger values of ψ cause the distribution to be more centred around and concentrated on the expected

value of 1. This is in line with (2.2.5) which shows that $Var(W) \to 0$ as $\psi \to \infty$. Figure 2.2 shows the same $UIG(\psi)$ densities on a log-scale, highlighting the differences in tail behaviour.

The fact that $E(W) = 1$ regardless of the value of $\psi > 0$ is the reason for calling this the "unit" IG distribution. If we let $V = \delta W / \gamma$ for any $\delta, \gamma > 0$ subject to $\sqrt{\delta\gamma} = \psi$, it follows from (1.6.4) that the density of V is

$$f_V(v) = \frac{\gamma}{\delta} f_W\left(\frac{\gamma v}{\delta}\right) = \left(\frac{\gamma}{\delta}\right)^{-1/2} \frac{v^{-3/2}}{2K_{1/2}(\delta\gamma)} \exp\left(-\frac{1}{2}\left(\delta^2 v^{-1} + \gamma^2 v\right)\right). \qquad (2.2.6)$$

Thus, comparing (2.2.6) with (2.2.1), if W is $UIG(\psi)$-distributed, then for any $\delta, \gamma > 0$ subject to $\sqrt{\delta\gamma} = \psi$, $V = \delta W / \gamma$ is $IG(\delta, \gamma)$-distributed.

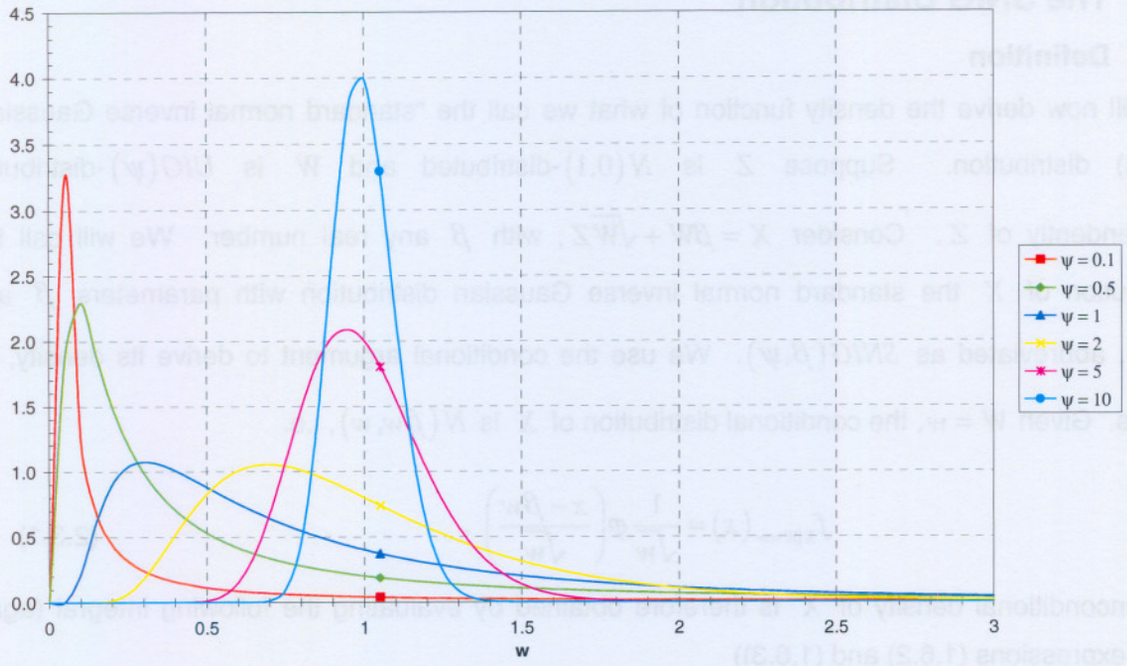

Figure 2.1: $UIG(\psi)$ -densities.

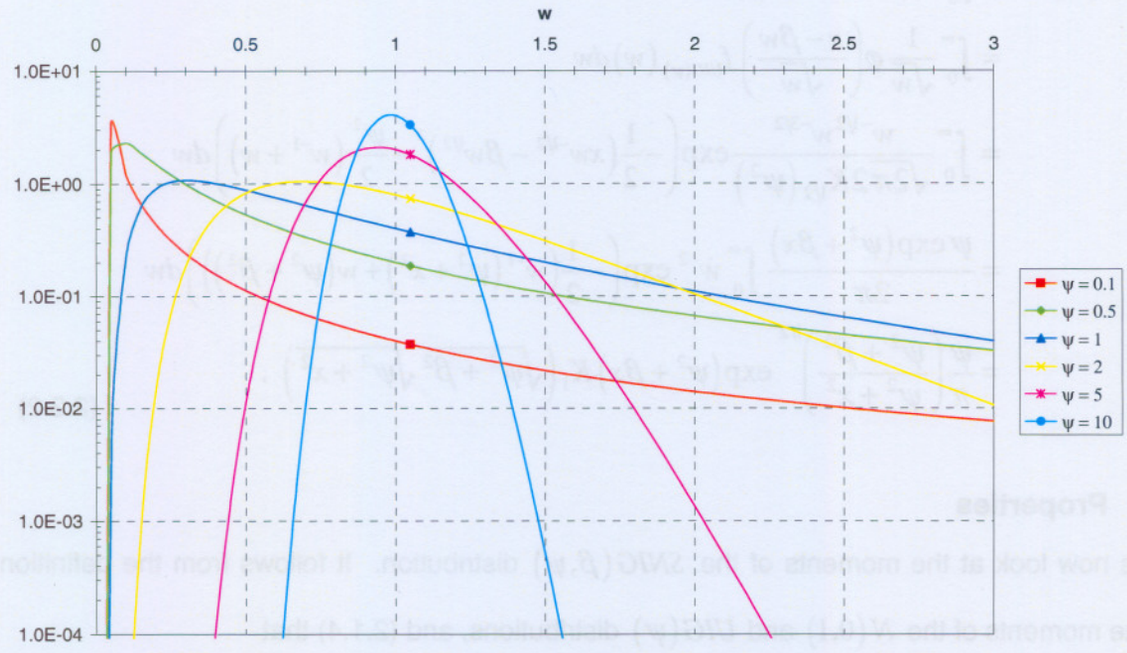

Figure 2.2: $UIG(\psi)$ -densities on a log-scale.

2.3 The SNIG Distribution

2.3.1 Definition

We will now derive the density function of what we call the "standard normal inverse Gaussian" (SNIG) distribution. Suppose Z is $N(0,1)$-distributed and W is $UIG(\psi)$-distributed independently of Z. Consider $X = \beta W + \sqrt{W}Z$, with β any real number. We will call the distribution of X the standard normal inverse Gaussian distribution with parameters β and $\psi > 0$, abbreviated as $SNIG(\beta,\psi)$. We use the conditional argument to derive its density, as follows. Given $W = w$, the conditional distribution of X is $N(\beta w, w)$, i.e.

$$f_{X|W=w}(x) = \frac{1}{\sqrt{w}}\varphi\left(\frac{x-\beta w}{\sqrt{w}}\right). \tag{2.3.1}$$

The unconditional density of X is therefore obtained by evaluating the following integral (again using expressions (1.6.2) and (1.6.3))

$$\begin{aligned}
f_{SNIG(\beta,\psi)}(x) &= \int_0^\infty f_{X|W=w}(x) f_{UIG(\psi)}(w)\,dw \\
&= \int_0^\infty \frac{1}{\sqrt{w}}\varphi\left(\frac{x-\beta w}{\sqrt{w}}\right) f_{UIG(\psi)}(w)\,dw \\
&= \int_0^\infty \frac{w^{-1/2}w^{-3/2}}{\sqrt{2\pi}\,2K_{1/2}(\psi^2)} \exp\left(-\frac{1}{2}\left(xw^{-1/2} - \beta w^{1/2}\right)^2 - \frac{\psi^2}{2}\left(w^{-1}+w\right)\right) dw \\
&= \frac{\psi \exp(\psi^2 + \beta x)}{2\pi} \int_0^\infty w^{-2} \exp\left(-\frac{1}{2}\left(w^{-1}\left(\psi^2 + x^2\right) + w\left(\psi^2 + \beta^2\right)\right)\right) dw \\
&= \frac{\psi}{\pi}\left(\frac{\psi^2+\beta^2}{\psi^2+x^2}\right)^{1/2} \exp(\psi^2 + \beta x) K_1\left(\sqrt{\psi^2+\beta^2}\sqrt{\psi^2+x^2}\right).
\end{aligned} \tag{2.3.2}$$

2.3.2 Properties

Let us now look at the moments of the $SNIG(\beta,\psi)$ distribution. It follows from the definition of X, the moments of the $N(0,1)$ and $UIG(\psi)$ distributions, and (2.1.4) that

$$
\begin{aligned}
E(X) &= \beta E(W) + E\left(\sqrt{W}\right)E(Z) = \beta \\
E\left(X^2\right) &= \beta^2 E\left(W^2\right) + 2\beta E\left(W^{3/2}\right)E(Z) + E(W)E\left(Z^2\right) = \beta^2 E\left(W^2\right) + 1 \\
&= \beta^2\left(1+\frac{1}{\psi^2}\right)+1 \\
E\left(X^3\right) &= \beta^3 E\left(W^3\right) + 3\beta^2 E\left(W^{5/2}Z\right) + 3\beta E\left(W^2Z^2\right) + E\left(W^{3/2}Z^3\right) \\
&= \beta^3\left(1+\frac{3}{\psi^2}+\frac{3}{\psi^4}\right)+3\beta\left(1+\frac{1}{\psi^2}\right) \\
E\left(X^4\right) &= \beta^4 E\left(W^4\right) + 4\beta^3 E\left(W^{7/2}Z\right) + 6\beta^2 E\left(W^3Z^2\right) + 4\beta E\left(W^{5/2}Z^3\right) + E\left(W^2Z^4\right) \\
&= \beta^4\left(1+\frac{6}{\psi^2}+\frac{15}{\psi^4}+\frac{15}{\psi^6}\right)+6\beta^2\left(1+\frac{3}{\psi^2}+\frac{3}{\psi^2}\right)+3\left(1+\frac{1}{\psi^2}\right)
\end{aligned}
\tag{2.3.3}
$$

so that

$$
\begin{aligned}
Var(X) &= \beta^2 Var(W) = 1+\beta^2/\psi^2 \\
Skew(X) &= \frac{E\left(X^3\right)-3E(X)E\left(X^2\right)+2\left(E(X)\right)^3}{Var(X)^{3/2}} \\
&= \frac{\left(3\beta/\psi^2\right)\left(1+\beta^2/\psi^2\right)}{\left(1+\beta^2/\psi^2\right)^{3/2}} \\
&= \frac{3\beta}{\sqrt{\psi^4+\beta^2\psi^2}} \\
Kurt(X) &= \frac{E\left(X^4\right)-4E(X)E\left(X^3\right)+6\left(E(X)\right)^2E\left(X^2\right)-3\left(E(X)\right)^4}{Var(X)^2}-3 \\
&= \frac{3/\psi^2+18\beta^2/\psi^4+15\beta^4/\psi^6}{\left(1+\beta^2/\psi^2\right)^2}\ .
\end{aligned}
\tag{2.3.4}
$$

When $\beta = 0$ we have $E(X) = 0$ and $Var(X) = 1$, and calling the distribution of X the "standard" NIG distribution seems particularly appropriate. We will, however, also use the name SNIG for the case $\beta \neq 0$.

Figure 2.3 shows $SNIG(-1,\psi)$ densities for several values of ψ while Figure 2.4 shows these densities on a log-scale. From (2.3.2) it can easily be seen that $f_{SNIG(-\beta,\psi)}(x) = f_{SNIG(\beta,\psi)}(-x)$. This means that the density of $SNIG(\beta,\psi)$ can be obtained from the density of $SNIG(-\beta,\psi)$ by flipping the graph around the $x = 0$ axis. We see that smaller ψ-values cause the densities to be

less spread out with heavier tails, while the densities are more skewed to the left for $\beta < 0$ (and from symmetry more skewed to the right for $\beta > 0$).

Figure 2.5 and Figure 2.6 with log-scale display $SNIG(\beta, \psi)$ densities for four (β, ψ) pairs. We can again see how smaller ψ-values cause the distributions to be more peaked with heavier tails and, with $\beta > 0$, more skewed to the right. It is also clear that $\beta = 0$ is the symmetric case while $\beta \neq 0$ affects location, scale, skewness and kurtosis (see (2.3.4)). Note that the $SNIG(-1, \psi)$ density with $\psi = 5$ is already quite close to the limiting $SNIG(-1, \infty) \equiv N(-1,1)$ density.

2.3.3 The SNIG as a Normal Mean-Variance Mixture

If $Y = \beta + \sqrt{W}Z$, with $Z \sim N(0,1)$ and W a positive valued random variable independent of Z, then Y is said to have a normal variance mixture distribution. As a special case, if $\beta = 0$ and W has an inverse Gamma distribution with parameters $\nu/2$ and $\nu/2$ (or, equivalently, ν/W is χ^2_ν-distributed), then Y has a Student's t-distribution with ν degrees of freedom (see Andrews and Mallows (1974)).

Furthermore, if $X = \beta W + \sqrt{W}Z$ then the density of X will be a normal mean-variance mixture. Other authors have studied such mean-variance mixtures, for example Barndorff-Nielsen (1977, 1978) obtains the generalised hyperbolic distribution as a mixture of normal distributions, where the mixing distribution is GIG. From the first line of (2.3.2) we see that the $SNIG(\beta, \psi)$ distribution is a mean-variance mixture of the normal distribution with $UIG(\psi)$ as mixing distribution.

Frey and McNeil (2003) generalise normal mean-variance mixtures by letting

$$X = g(W) + \sqrt{W}Z, \tag{2.3.5}$$

with $g : \mathbb{R}^+ \to \mathbb{R}$ and also extend (2.3.5) to multivariate normal mean-variance mixtures. We will limit ourselves to the univariate case and the choice above leading to the SNIG distribution.

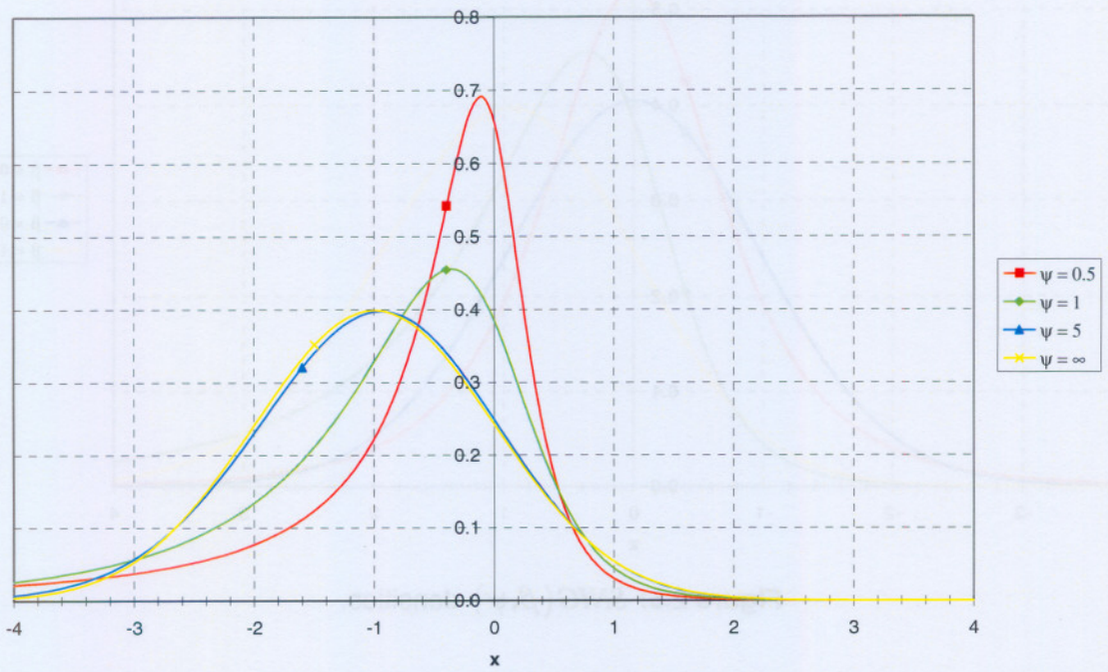

Figure 2.3: $SNIG(-1,\psi)$-densities.

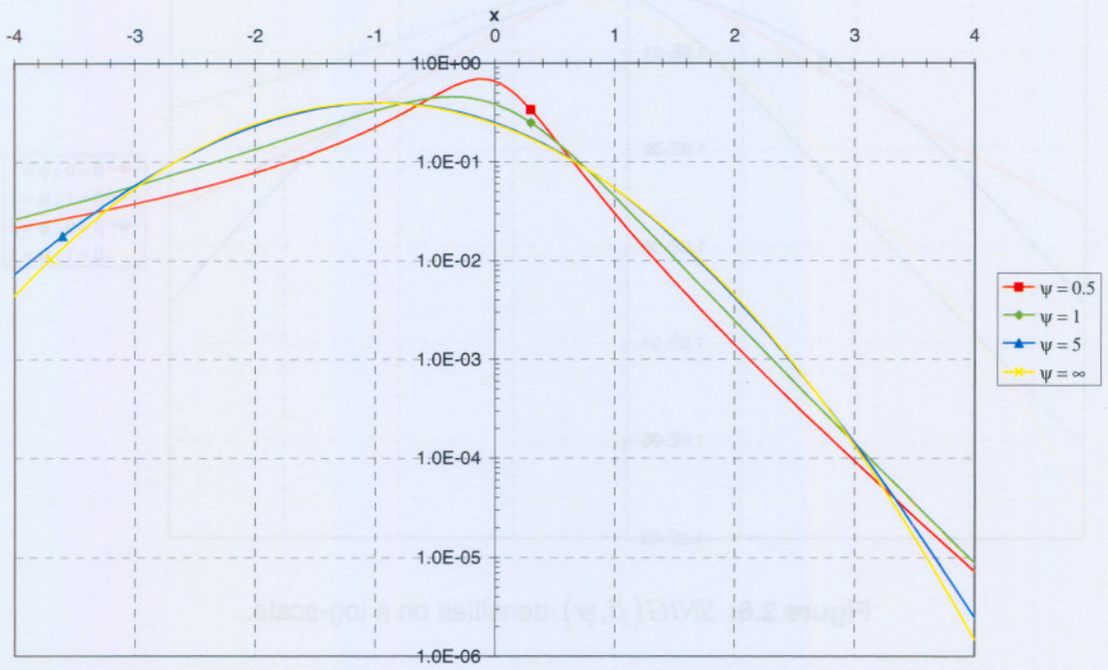

Figure 2.4: $SNIG(-1,\psi)$-densities on a log-scale.

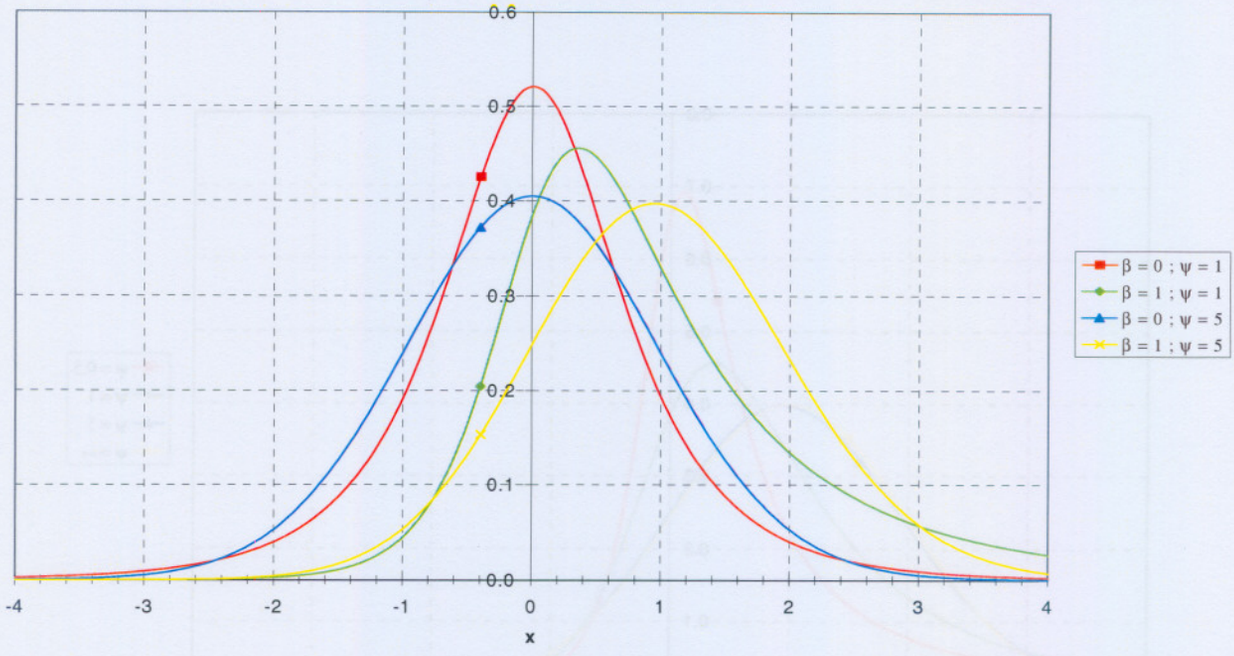

Figure 2.5: $SNIG(\beta,\psi)$-densities.

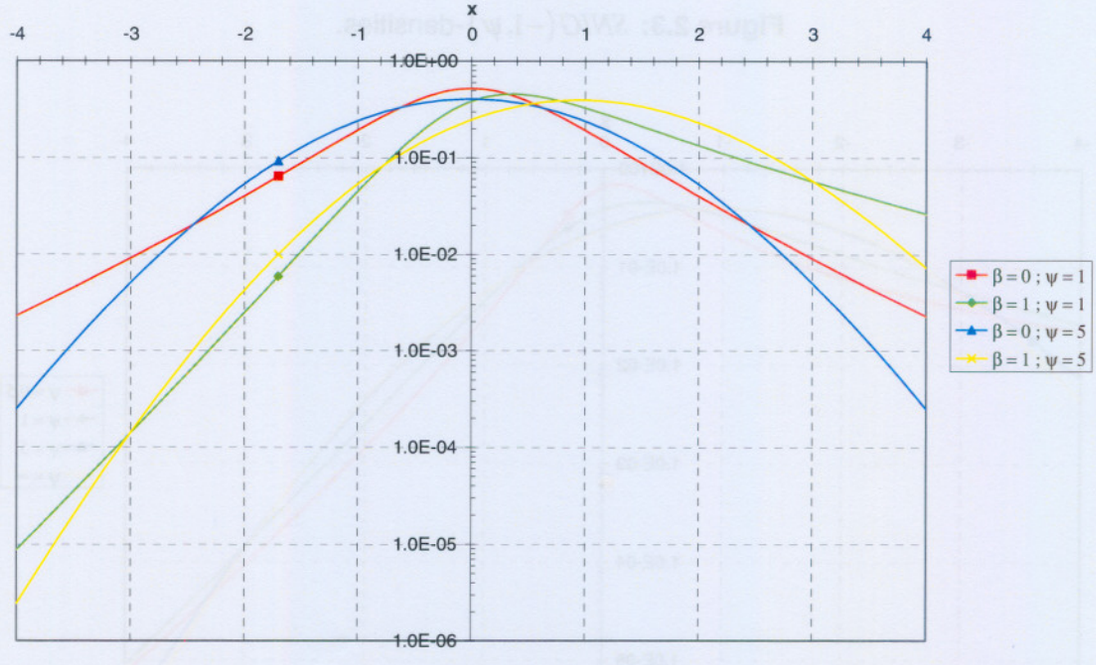

Figure 2.6: $SNIG(\beta,\psi)$-densities on a log-scale.

2.3.4 Relating the SNIG to the NIG

We will now see how making additional scale and translation transformations relates the SNIG distribution to the common parameterisation of the NIG distribution, as used for example by Barndorff-Nielsen and Prause (2001) and Lillestøl (2000). In their parameterisation, the NIG density function with parameters $(\alpha_1, \beta_1, \mu_1, \delta_1)$ is given by

$$f_{NIG(\alpha_1,\beta_1,\mu_1,\delta_1)}(x) = \frac{\alpha_1 \exp\left(\varsigma + \beta_1 (x-\mu_1)\right) K_1\left(\alpha_1 \delta_1 q\left(\frac{x-\mu_1}{\delta_1}\right)\right)}{\pi q\left(\frac{x-\mu_1}{\delta_1}\right)} \tag{2.3.6}$$

where $0 < |\beta_1| < \alpha_1$, $-\infty < \mu_1 < \infty$, $\delta_1 > 0$, $\varsigma = \delta_1\sqrt{\alpha_1^2 - \beta_1^2}$ and $q(x) = \sqrt{1+x^2}$. α_1, β_1, μ_1 and δ_1 are steepness, asymmetry, location and scale parameters respectively.

Consider $Y = \mu + \delta X$ with X $SNIG(\beta, \psi)$-distributed as in Section 2.3.1, μ any real number and $\delta > 0$. The moments of Y follow immediately from the definition of Y and the moments of X, namely

$$\begin{aligned} E(Y) &= \mu + \delta E(X) = \mu + \delta\beta \\ Var(Y) &= \delta^2 Var(X) = \frac{\delta^2\beta^2}{\psi^2} + \delta^2 \end{aligned} \tag{2.3.7}$$

We can use (1.5.4) to get the density of Y as

$$\begin{aligned} f_Y(y) &= \frac{1}{\delta} f_X\left(\frac{y-\mu}{\delta}\right) \\ &= \frac{\psi}{\delta\pi}\left(\frac{\psi^2+\beta^2}{\psi^2+\left(\frac{y-\mu}{\delta}\right)^2}\right)^{1/2} \exp\left(\psi^2 + \beta\left(\frac{y-\mu}{\delta}\right)\right) K_1\left(\sqrt{\psi^2+\beta^2}\sqrt{\psi^2+\left(\frac{y-\mu}{\delta}\right)^2}\right) \\ &= \frac{1}{\pi\delta}\left(\frac{\psi^2+\beta^2}{1+\left(\frac{y-\mu}{\psi\delta}\right)}\right)^{1/2} \exp\left(\psi^2 + \beta\left(\frac{y-\mu}{\delta}\right)\right) K_1\left(\psi\sqrt{\psi^2+\beta^2}\sqrt{1+\left(\frac{y-\mu}{\psi\delta}\right)^2}\right). \end{aligned} \tag{2.3.8}$$

If we now let $(\alpha_1, \beta_1, \mu_1, \delta_1) = \left(\sqrt{\psi^2+\beta^2}/\delta, \beta/\delta, \mu, \psi\delta\right)$ (2.3.6) becomes

$$f_{NIG(\alpha_1,\beta_1,\mu_1,\delta_1)}(x)=\frac{1}{\pi\delta}\left(\frac{\psi^2+\beta^2}{1+\left(\frac{x-\mu}{\psi\delta}\right)^2}\right)^{\frac{1}{2}}\exp\left(\psi^2+\beta\left(\frac{x-\mu}{\delta}\right)\right)K_1\left(\psi\sqrt{\psi^2+\beta^2}\sqrt{1+\left(\frac{x-\mu}{\psi\delta}\right)^2}\right). \quad (2.3.9)$$

Thus, comparing (2.3.8) and (2.3.9) we see that, if X is $SNIG(\beta,\psi)$-distributed, $Y=\mu+\delta X$ is $NIG\left(\sqrt{\psi^2+\beta^2}/\delta,\beta/\delta,\mu,\psi\delta\right)$-distributed.

As a special case, if we let $\mu=0$, $\delta=1$ and therefore

$$(\alpha_1,\beta_1,\mu_1,\delta_1)=\left(\sqrt{\psi^2+\beta^2},\beta,0,\psi\right) \quad (2.3.10)$$

we see that the $SNIG(\beta,\psi)$ distribution corresponds to the $NIG\left(\sqrt{\psi^2+\beta^2},\beta,0,\psi\right)$ distribution, since

$$\begin{aligned} f_{NIG\left(\sqrt{\psi^2+\beta^2},\beta,0,\psi\right)}(x) &= \frac{\psi}{\pi}\left(\frac{\psi^2+\beta^2}{\psi^2+x^2}\right)^{1/2}\exp\left(\psi^2+\beta x\right)K_1\left(\sqrt{\psi^2+\beta^2}\sqrt{\psi^2+x^2}\right) \\ &= f_{SNIG(\beta,\psi)}(x). \end{aligned} \quad (2.3.11)$$

Alternative Shape Parameters

Barndorff-Nielsen et al. (1985) also introduces two alternative steepness and asymmetry parameters for the NIG distribution, namely

$$\xi=\left(1+\delta_1\sqrt{\alpha_1^2-\beta_1^2}\right)^{-1/2} \text{ and } \chi=\frac{\beta_1\xi}{\alpha_1}. \quad (2.3.12)$$

The parameter space for these shape parameters is given by the so-called "shape triangle"

$$0\le|\chi|<\xi<1. \quad (2.3.13)$$

For the $SNIG(\beta,\psi)$ distribution as defined above, it follows from (2.3.10) and (2.3.12) that the steepness and asymmetry parameters ξ and χ are given by

$$\xi=\left(1+\psi^2\right)^{-1/2} \text{ and } \chi=\beta\xi/\sqrt{\psi^2+\beta^2}. \quad (2.3.14)$$

(2.3.14) may be inverted to obtain (β,ψ) for any given pair (ξ,χ) in $0\le|\chi|<\xi$. This yields

$$\psi=\left(\xi^{-2}-1\right)^{1/2} \text{ and } \beta=\psi\chi/\sqrt{\xi^2-\chi^2}. \quad (2.3.15)$$

The pair (β,ψ) has a simpler parameter space than the pair (ξ,χ), namely $-\infty<\beta<\infty,\psi>0$. Figure 2.7 gives an indication of the relation between these two parameter pairs using a few corresponding points. For instance, the square shows that $(\xi,\chi)=(0.2,0.1)$ corresponds to $(\beta,\psi)\approx(0.015,4.9)$. We shall use the (β,ψ) rather than the (ξ,χ) parameterisation throughout this thesis.

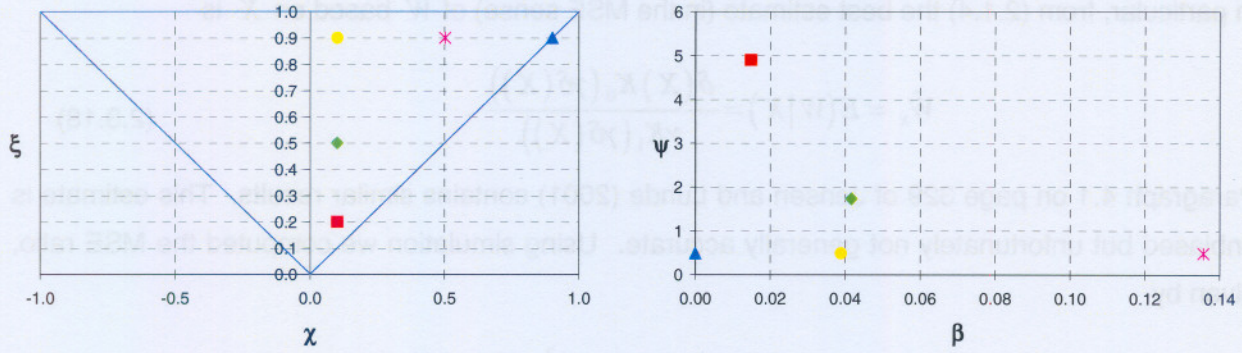

Figure 2.7: Relation between (ξ,χ) and (β,ψ) NIG parameter pairs.

2.3.5 Estimating W and Z from X

In Section 3.3 we shall find ourselves in situations where we have observed X, which can be modelled as having the distribution of $\beta W+\sqrt{W}Z$ with W and Z latent (unobserved) variables. In such situations we shall want to estimate W (and Z) from an observation on $X=\beta W+\sqrt{W}Z$. The minimum mean squared error (MSE) estimator of W is $\hat{W}_X=E(W|X)$ and we need the conditional distribution of W given X to calculate it. For this purpose let

$$\delta(x)=\sqrt{\psi^2+x^2} \text{ and } \gamma=\sqrt{\psi^2+\beta^2}\ . \tag{2.3.16}$$

It then follows that (see (2.3.2))

$$\begin{aligned}
&\frac{1}{\sqrt{w}}\varphi\left(\frac{x-\beta w}{\sqrt{w}}\right)f_{UIG(\psi)}(w)\\
&=\frac{\psi w^{-2}}{2\pi}\exp\left(\psi^2+\beta x\right)\exp\left(-\frac{1}{2}\left(\left(\psi^2+x^2\right)w^{-1}+\left(\psi^2+\beta^2\right)w\right)\right)\\
&=\frac{\psi}{\pi}\left(\frac{\gamma^2}{\delta^2(x)}\right)^{\frac{1}{2}}\exp\left(\psi^2+\beta x\right)K_1\left(\gamma\delta(x)\right)\left(\frac{\gamma}{\delta(x)}\right)^{-1}\frac{w^{-2}}{2K_1\left(\gamma\delta(x)\right)}\exp\left(-\frac{1}{2}\left(\frac{\delta^2(x)}{w}+\gamma w\right)\right)\\
&=f_{SNIG(\beta,\psi)}(x)\,f_{GIG(-1,\delta(x),\gamma)}(w)\ .
\end{aligned} \tag{2.3.17}$$

Several facts follow from (2.3.17). With X, Z and W as above, the density of X given $W=w$ is $\varphi\left((x-\beta w)/\sqrt{w}\right)/\sqrt{w}$ so that the first line, and therefore the last line, of (2.3.17) represents the joint density of X and W. Integrating the last line of (2.3.17) over w, we see again that the marginal of X is $f_{SNIG(\beta,\psi)}(x)$, thus rederiving (2.3.2). Moreover, dividing the last line of (2.3.17) by $f_{SNIG(\beta,\psi)}(x)$, we see that the conditional distribution of W given $X=x$ is $GIG\left(-1,\delta(x),\gamma\right)$. In particular, from (2.1.4) the best estimate (in the MSE sense) of W based on X is

$$\hat{W}_X = E\left(W|X\right) = \frac{\delta(X)K_0\left(\gamma\delta(X)\right)}{\gamma K_1\left(\gamma\delta(X)\right)}\ . \tag{2.3.18}$$

Paragraph 4.1 on page 329 of Jensen and Lunde (2001) contains similar results. This estimate is unbiased but unfortunately not generally accurate. Using simulation we computed the MSE ratio, given by

$$\text{MSE Ratio} = \frac{E\left(\hat{W}_X - W\right)^2}{Var(W)}\ , \tag{2.3.19}$$

corresponding to this estimate and show this in Table 2.1 (we get to the entry for $\hat{Z}_X$ in a moment). The MSE ratios for $\hat{W}_X$ are smaller for smaller values of ψ and values of β further away from 0. The cases of β close to 0 and ψ in the vicinity of 1 are of most practical interest and for these the MSE ratios are quite large, so that X explains only a smallish fraction of the variability of W. To shed some light on why this is so, consider (2.3.16), (2.3.18) and Figure 2.8. It is clear from (2.3.16) that $\delta(X)$ (and therefore $\delta(X)/\gamma$ and $\gamma\delta(X)$) is monotone increasing in X. Furthermore, Figure 2.8 shows that the ratio $K_0(z)/K_1(z)$ is monotone increasing in z. Consequently, (2.3.18) reaches a minimum, which is the smallest possible estimate $\hat{W}_X$ for W, in the argument $X=0$, i.e. $\hat{W}_X \geq \psi K_0(\gamma\psi)/\gamma K_1(\gamma\psi)$ since $\delta(0)=\psi$ from (2.3.16). This minimum is given in Table 2.2 as a function of β and ψ. The minimum possible estimate $\hat{W}_X$ for W when, say, $\beta=0$ and $\psi=2$ is 0.894, while W (from a $UIG(2)$ distribution) frequently takes on values much closer to 0, as can be seen from Figure 2.1. This is one reason why the MSE ratios are so large, especially for small β and large ψ. Nevertheless, $\hat{W}_X$ is the best estimator of W in the MSE sense.

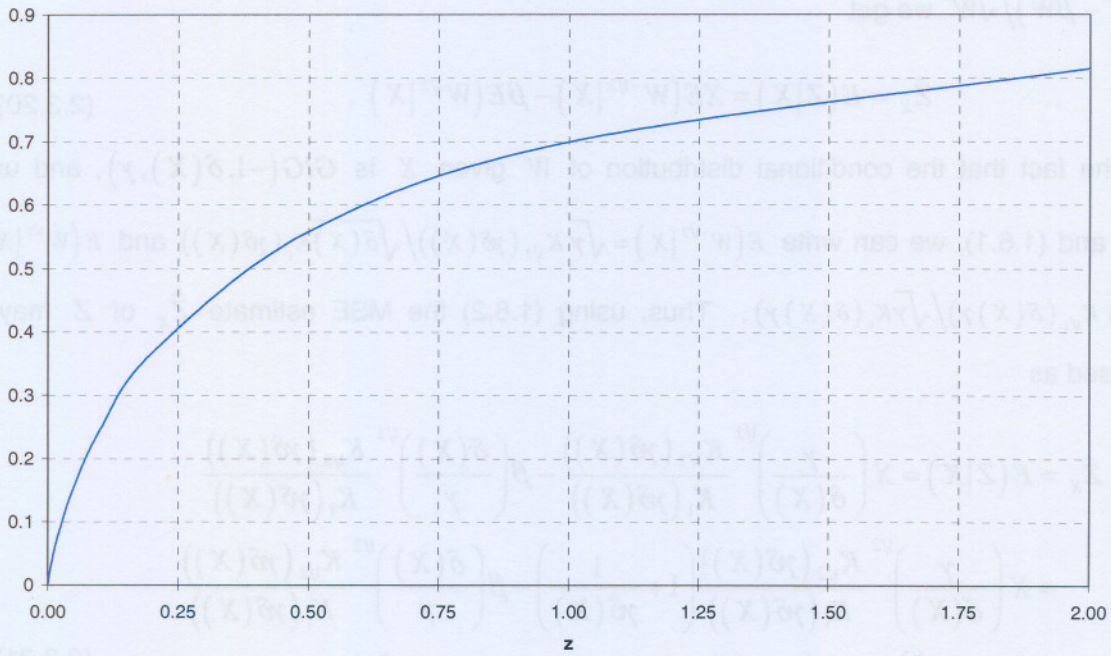

Figure 2.8: $K_0(z)/K_1(z)$ is monotone increasing.

	β									
	0		0.25		0.5		1		2	
ψ	$\hat{W}_X$	$\hat{Z}_X$	$\hat{W}_X$	$\hat{Z}_X$	$\hat{W}_X$	$\hat{Z}_X$	$\hat{W}_X$	$\hat{Z}_X$	$\hat{W}_X$	$\hat{Z}_X$
0.10	0.566	0.211	0.114	0.228	0.032	0.251	0.010	0.293	0.002	0.364
0.25	0.584	0.195	0.345	0.224	0.158	0.273	0.051	0.363	0.014	0.498
0.50	0.653	0.158	0.543	0.190	0.375	0.256	0.171	0.398	0.055	0.593
1.00	0.780	0.098	0.737	0.121	0.643	0.180	0.431	0.338	0.186	0.595
2.00	0.907	0.043	0.894	0.054	0.858	0.084	0.738	0.185	0.474	0.427
5.00	0.981	0.009	0.979	0.012	0.972	0.018	0.945	0.045	0.850	0.138

Table 2.1: MSE ratios for $\hat{W}_X$ and $\hat{Z}_X$ given X.

	β				
ψ	0.00	0.25	1.00	2.00	5.00
0.25	0.182	0.161	0.101	0.069	0.037
0.50	0.411	0.388	0.260	0.171	0.084
1.00	0.699	0.684	0.538	0.371	0.179
2.00	0.894	0.888	0.808	0.652	0.355
5.00	0.981	0.979	0.962	0.912	0.697

Table 2.2: Lower bounds for $\hat{W}_X$ for different values of β and ψ.

We may also want to estimate Z from an observation on X. From the expression $Z=(X-\beta W)/\sqrt{W}$ we get

$$\hat{Z}_X = E(Z|X) = XE(W^{-1/2}|X) - \beta E(W^{1/2}|X)\ . \tag{2.3.20}$$

From the fact that the conditional distribution of W given X is $GIG(-1,\delta(X),\gamma)$, and using (2.1.4) and (1.6.1), we can write $E(W^{-1/2}|X) = \sqrt{\gamma}\,K_{3/2}(\gamma\delta(X))/\sqrt{\delta(X)}K_1(\gamma\delta(X))$ and $E(W^{1/2}|X) = \sqrt{\delta(X)}\,K_{3/2}(\delta(X)\gamma)/\sqrt{\gamma}K_1(\delta(X)\gamma)$. Thus, using (1.6.2) the MSE estimate $\hat{Z}_X$ of Z may be expressed as

$$\begin{aligned}
\hat{Z}_X = E(Z|X) &= X\left(\frac{\gamma}{\delta(X)}\right)^{1/2}\frac{K_{3/2}(\gamma\delta(X))}{K_1(\gamma\delta(X))} - \beta\left(\frac{\delta(X)}{\gamma}\right)^{1/2}\frac{K_{1/2}(\gamma\delta(X))}{K_1(\gamma\delta(X))} \\
&= X\left(\frac{\gamma}{\delta(X)}\right)^{1/2}\frac{K_{1/2}(\gamma\delta(X))}{K_1(\gamma\delta(X))}\left(1+\frac{1}{\gamma\delta(X)}\right) - \beta\left(\frac{\delta(X)}{\gamma}\right)^{1/2}\frac{K_{1/2}(\gamma\delta(X))}{K_1(\gamma\delta(X))} \\
&= \left(\frac{\gamma}{\delta(X)}\right)^{1/2}\frac{K_{1/2}(\gamma\delta(X))}{K_1(\gamma\delta(X))}\left(X\left(1+\frac{1}{\gamma\delta(X)}\right) - \beta\frac{\delta(X)}{\gamma}\right).
\end{aligned} \tag{2.3.21}$$

Table 2.1 also shows the corresponding MSE ratios for $\hat{Z}_X$. Here $\hat{Z}_X$ tends to estimate Z more accurately than $\hat{W}_X$ estimates W. In contrast to the MSE ratios for $\hat{W}_X$, the MSE ratios for $\hat{Z}_X$ are smaller for larger values of ψ and values of β closer to 0. Later on we may use $\hat{W}_X$ and $\hat{Z}_X$ for model checking purposes in which case their distributions are required. If $\hat{W}_X$ and $\hat{Z}_X$ were very accurate, their distributions would have been $UIG(\psi)$ and $N(0,1)$ respectively, but we cannot rely on this. Their exact distributions seem quite hard to compute theoretically but in practice we can generate enough realisations of $\hat{W}_X$ and $\hat{Z}_X$ at given β and ψ to obtain sufficiently accurate empirical estimates of these distributions.

This generation process involves the following steps:

- Generate $Z \sim N(0,1)$ and $W \sim UIG(\psi)$ (see Section 2.8).
- Set $X = \beta W + \sqrt{W}Z$.
- Compute $\hat{W}_X$ and $\hat{Z}_X$ from (2.3.18) and (2.3.21).
- Repeat as many times as required.

2.4 The SCIG Distribution

2.4.1 Definition

We will now define what we call the "standard chi-squared inverse Gaussian" (SCIG) distribution. Let W be $UIG(\psi)$-distributed as in Section 2.3. Suppose that C is χ^2_ι-distributed with $\iota > 0$, independently of W. Thus

$$f_{\chi^2_\iota}(c) = \frac{c^{\iota/2-1}\exp(-c/2)}{\Gamma(\iota/2)2^{\iota/2}},\ c>0\ . \tag{2.4.1}$$

Consider the distribution of $V = WC$, which may be called the standard chi-squared inverse Gaussian distribution (abbreviated $SCIG(\psi,\iota)$). From (1.5.4) it follows that the conditional distribution of V given $W = w$ is given by

$$f_{V|W=w}(v) = \frac{1}{w} f_{\chi^2_\iota}\left(\frac{v}{w}\right). \tag{2.4.2}$$

Multiplying (2.4.2) by the density of W and integrating out w, we get the density of V as

$$\begin{aligned} f_{SCIG(\psi,\iota)}(v) &= \int_0^\infty \frac{1}{w} f_{\chi^2_\iota}\left(\frac{v}{w}\right) f_{UIG(\psi)}(w)\,dw \\ &= \frac{\psi\exp(\psi^2)v^{\iota/2-1}}{\sqrt{\pi}\Gamma\left(\frac{\iota}{2}\right)2^{\frac{\iota+1}{2}}} \int_0^\infty w^{-\frac{\iota+3}{2}} \exp\left(-\frac{1}{2}\left(\frac{\psi^2+v}{w}+\psi^2 w\right)\right)dw \\ &= \frac{\psi^{\frac{\iota+3}{2}}\exp(\psi^2)}{\sqrt{\pi}\Gamma\left(\frac{\iota}{2}\right)2^{\frac{\iota-1}{2}}} \frac{v^{\iota/2-1}}{(\psi^2+v)^{\frac{\iota+1}{4}}} K_{\frac{\iota+1}{2}}\left(\psi\sqrt{\psi^2+v}\right). \end{aligned} \tag{2.4.3}$$

2.4.2 Properties

Moments of the $SCIG(\psi,\iota)$ distribution follow easily from the definitions and the moments of the $UIG(\psi)$ and χ^2_ι distributions, namely

$$\begin{aligned} E(V) &= E(W)E(C) = \iota \\ E(V^2) &= E(W^2)E(C^2) = \left(1+\frac{1}{\psi^2}\right)\iota(\iota+2) = \iota^2+2\iota+\frac{\iota(\iota+2)}{\psi^2} \end{aligned} \tag{2.4.4}$$

so that

$$Var(V) = E(V^2) - (E(V))^2 = 2\iota + \frac{\iota(\iota+2)}{\psi^2}. \qquad (2.4.5)$$

From the law of large numbers and the fact that C is χ^2_ι-distributed it follows that, if $\iota \to \infty$, the distribution of C/ι will converge to the degenerate distribution with all its probability on 1. Since $V/\iota = WC/\iota$, the distribution of V/ι will converge to the $UIG(\psi)$ distribution as $\iota \to \infty$. Figure 2.9 compares the distribution of V/ι with that of the $UIG(1)$ distribution for different values of ι. They compare very well, even for relatively small values of ι. For $\iota = 10$ the distributions of V/ι and $UIG(1)$ are almost identical.

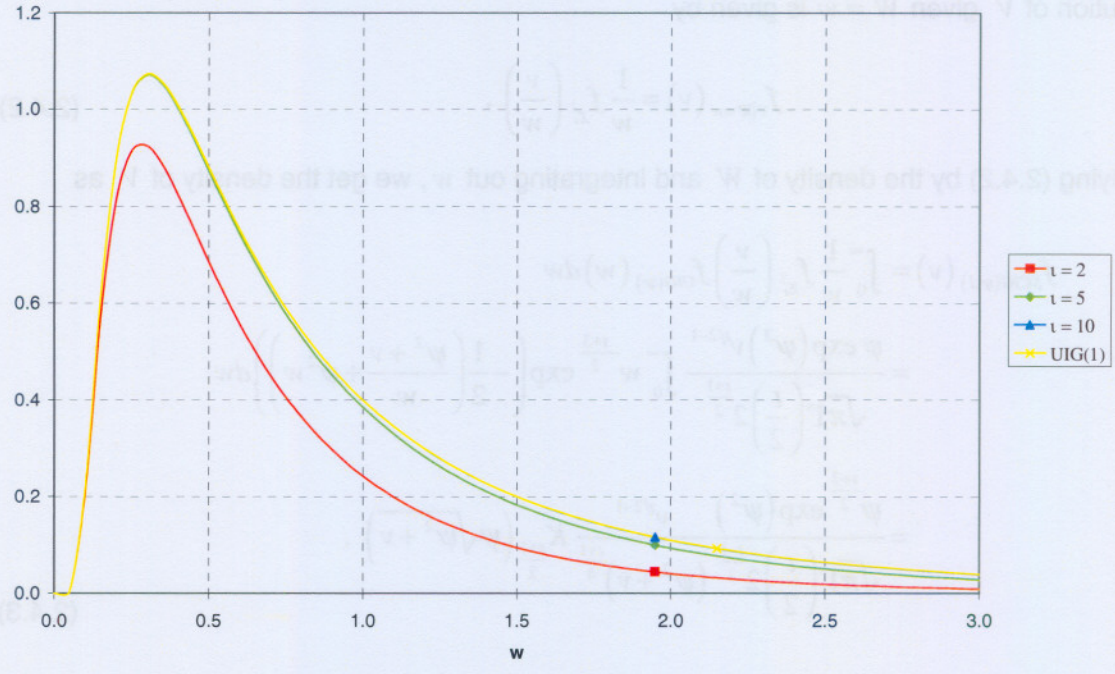

Figure 2.9: Densities of $SCIG(1,\iota)/\iota$ and $UIG(1)$ illustrating that $SCIG(1,\iota)/\iota \to UIG(1)$ as $\iota \to \infty$.

Figure 2.10 shows $SCIG(\psi,\iota)$-densities for different values of the parameters, while Figure 2.11 highlights the difference in tail behaviour. Smaller values of ψ cause the densities to be less spread out with heavier tails, while smaller ι-values lead to higher peaks.

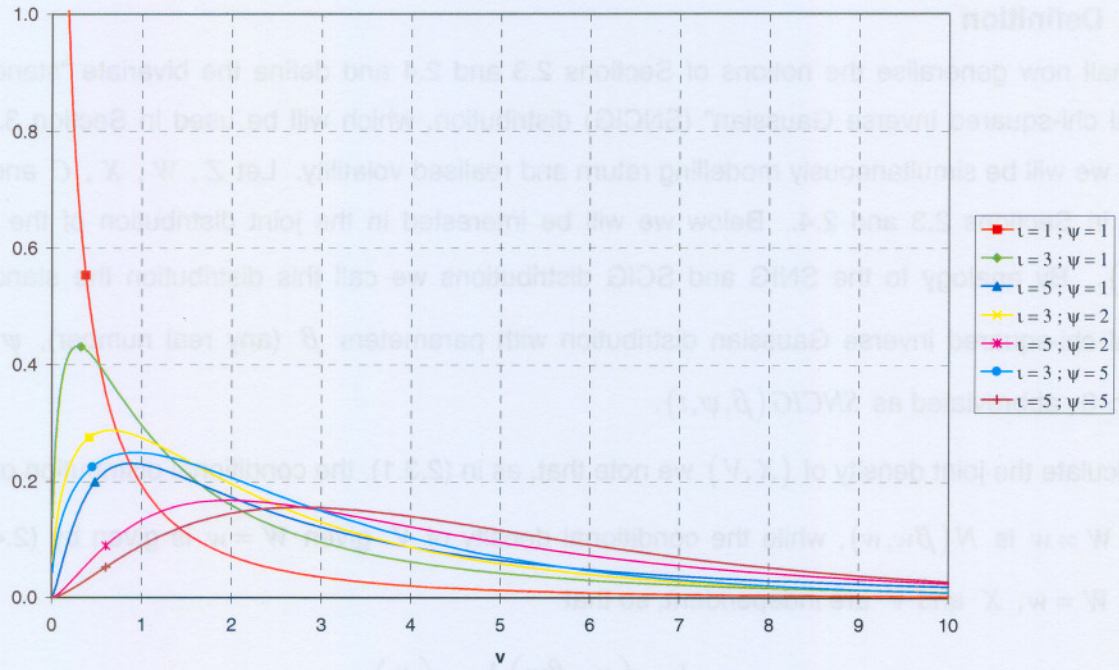

Figure 2.10: $SCIG(\psi,\iota)$-densities.

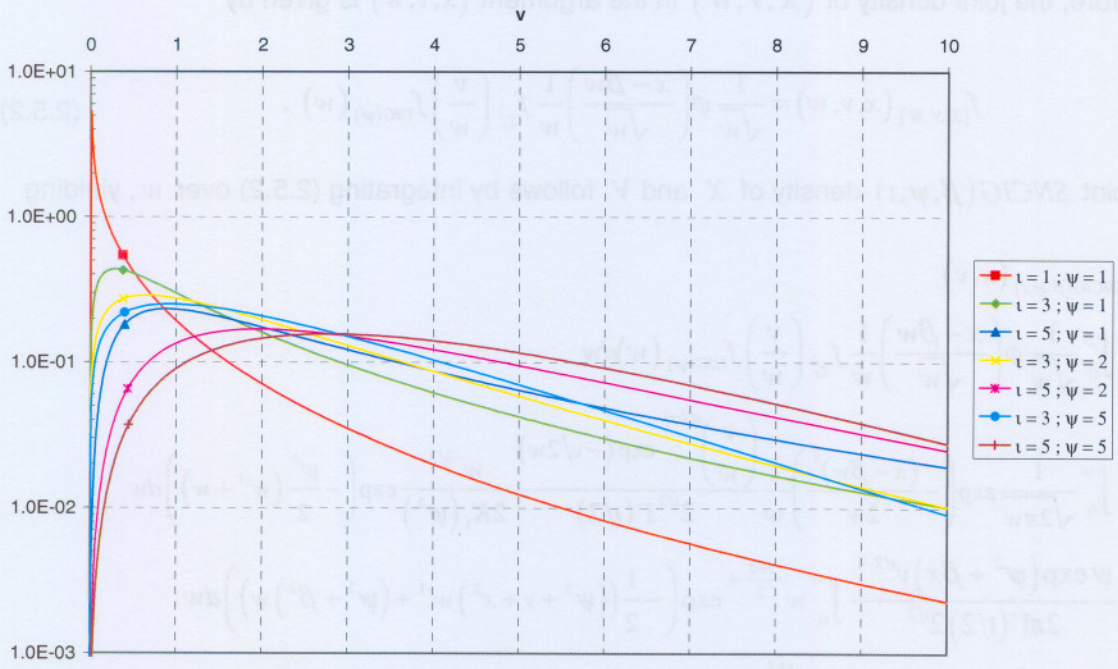

Figure 2.11: $SCIG(\psi,\iota)$-densities on a log-scale.

2.5 The SNCIG Distribution

2.5.1 Definition

We shall now generalise the notions of Sections 2.3 and 2.4 and define the bivariate "standard normal chi-squared inverse Gaussian" (SNCIG) distribution, which will be used in Section 3.2.3, where we will be simultaneously modelling return and realised volatility. Let Z, W, X, C and V be as in Sections 2.3 and 2.4. Below we will be interested in the joint distribution of the pair (X,V). By analogy to the SNIG and SCIG distributions we call this distribution the standard normal chi-squared inverse Gaussian distribution with parameters β (any real number), $\psi>0$ and $\iota>0$, abbreviated as $SNCIG(\beta,\psi,\iota)$.

To calculate the joint density of (X,V) we note that, as in (2.3.1), the conditional distribution of X given $W=w$ is $N(\beta w,w)$, while the conditional density of V given $W=w$ is given by (2.4.2). Given $W=w$, X and V are independent, so that

$$f_{(X,V)|W=w}(x,v)=\frac{1}{\sqrt{w}}\varphi\left(\frac{x-\beta w}{\sqrt{w}}\right)\frac{1}{w}f_{\chi^2_\iota}\left(\frac{v}{w}\right). \tag{2.5.1}$$

Therefore, the joint density of (X,V,W) in the argument (x,v,w) is given by

$$f_{(X,V,W)}(x,v,w)=\frac{1}{\sqrt{w}}\varphi\left(\frac{x-\beta w}{\sqrt{w}}\right)\frac{1}{w}f_{\chi^2_\iota}\left(\frac{v}{w}\right)f_{UIG(\psi)}(w). \tag{2.5.2}$$

The joint $SNCIG(\beta,\psi,\iota)$-density of X and V follows by integrating (2.5.2) over w, yielding

$$\begin{aligned}
&f_{SNCIG(\beta,\psi,\iota)}(x,v)\\
&=\int_0^\infty\frac{1}{\sqrt{w}}\varphi\left(\frac{x-\beta w}{\sqrt{w}}\right)\frac{1}{w}f_{\chi^2_\iota}\left(\frac{v}{w}\right)f_{UIG(\psi)}(w)dw\\
&=\int_0^\infty\frac{1}{\sqrt{2\pi w}}\exp\left(-\frac{(x-\beta w)^2}{2w}\right)\frac{1}{w}\frac{\left(\frac{v}{w}\right)^{\iota/2-1}\exp(-v/2w)}{2^{\iota/2}\Gamma(\iota/2)}\frac{w^{-3/2}}{2K_1(\psi^2)}\exp\left(-\frac{\psi^2}{2}\left(w^{-1}+w\right)\right)dw\\
&=\frac{\psi\exp(\psi^2+\beta x)v^{\iota/2-1}}{2\pi\Gamma(\iota/2)2^{\iota/2}}\int_0^\infty w^{-\frac{\iota+2}{2}-1}\exp\left(-\frac{1}{2}\left(\left(\psi^2+v+x^2\right)w^{-1}+\left(\psi^2+\beta^2\right)w\right)\right)dw\\
&=\frac{\psi}{\pi\Gamma(\iota/2)2^{\iota/2}}\left(\frac{\psi^2+\beta^2}{\psi^2+v+x^2}\right)^{\frac{\iota+2}{4}}v^{\iota/2-1}\exp(\psi^2+\beta x)K_{\frac{\iota}{2}+1}\left(\sqrt{\psi^2+\beta^2}\sqrt{\psi^2+v+x^2}\right).
\end{aligned} \tag{2.5.3}$$

2.5.2 Properties

Expectations and variances of X and V are given in Sections 2.3.2 and 2.4.2. The covariance between X and V is given by

$$\begin{aligned} Cov(X,V) &= E(XV) - E(X)E(V) \\ &= E\left(\beta CW^2\right) + E\left(CW^{3/2}Z\right) - E(X)E(V) \\ &= \beta\iota\left(1+\frac{1}{\psi^2}\right) - \beta\iota \\ &= \frac{\beta\iota}{\psi^2} \end{aligned} \tag{2.5.4}$$

It is interesting to note that in the symmetric case $\beta = 0$, $X = \sqrt{W}Z$ and $V = WC$ are uncorrelated even though they are not independent.

Figures 2.12 to 2.15 show $SNCIG(\beta,\psi,\iota)$-densities for different values of the parameters. Comparing Figures 2.12 and 2.13 we see how $\beta \neq 0$ shifts the density along the x-axis. Comparing Figure 2.13 with Figures 2.14 and 2.15 it is clear that larger ψ-values and larger ι-values shift the density along the positive v-axis and reduce the peakedness, causing the distribution to be more spread out.

2.5.3 Estimating W, Z and C from X and V

Analogous to Section 2.3.5, we shall find ourselves in situations in Section 3.3 where we have observed X and V, which can be modelled as having the joint distribution of $\beta W + \sqrt{W}Z$ and WC, with W, Z and C latent variables. In such situations we want to estimate W, Z and C from observations on $X = \beta W + \sqrt{W}Z$ and $V = WC$.

The minimum MSE estimator of W is $\hat{W}_{XV} = E(W|X,V)$. To calculate this we need the conditional distribution of W given $(X,V) = (x,v)$. For this purpose, we note the identity

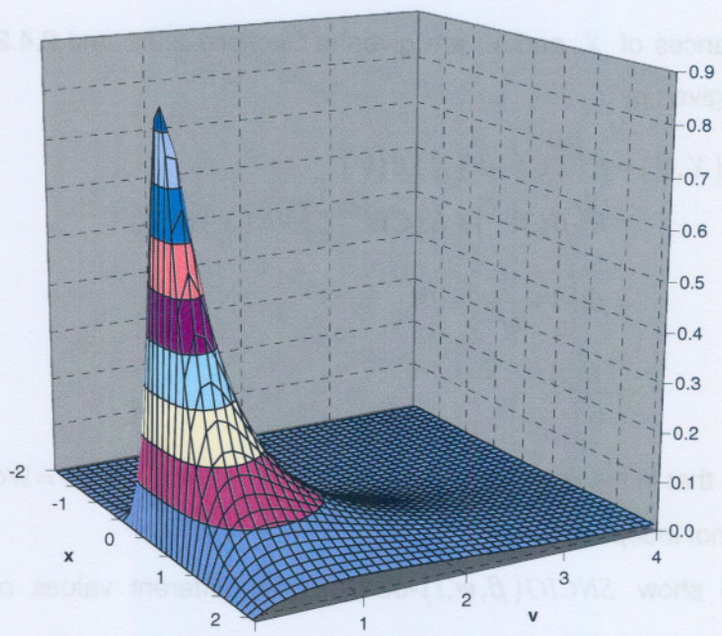

Figure 2.12: The $SNCIG(3, 0.5, 4)$-density.

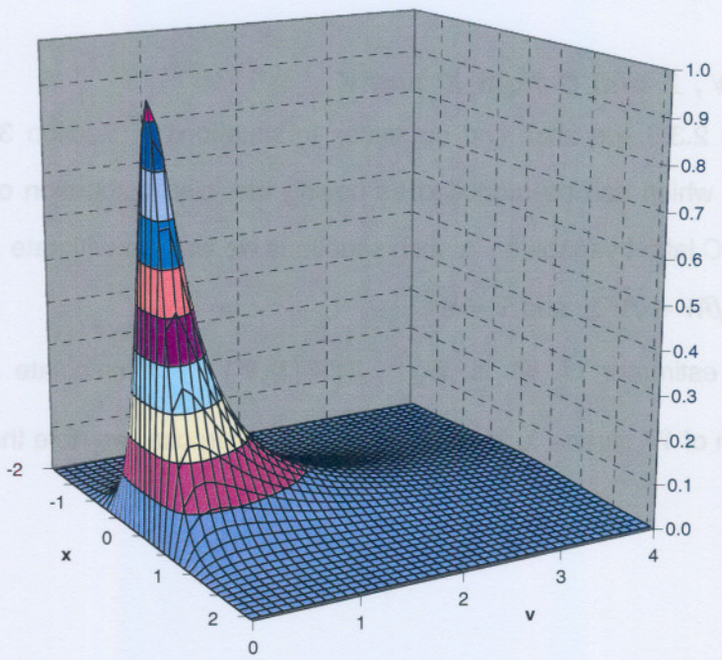

Figure 2.13: The $SNCIG(0, 0.5, 4)$-density.

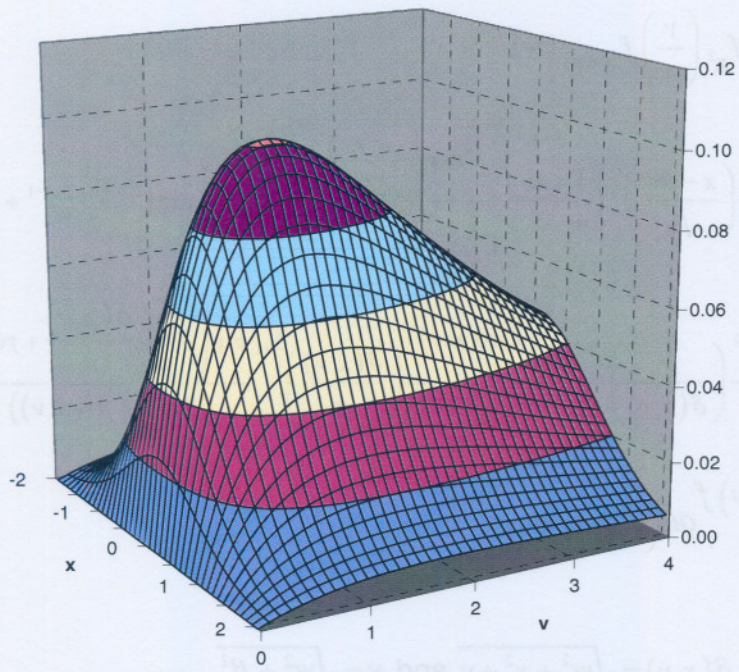

Figure 2.14: The $SNCIG(0,2,4)$-density.

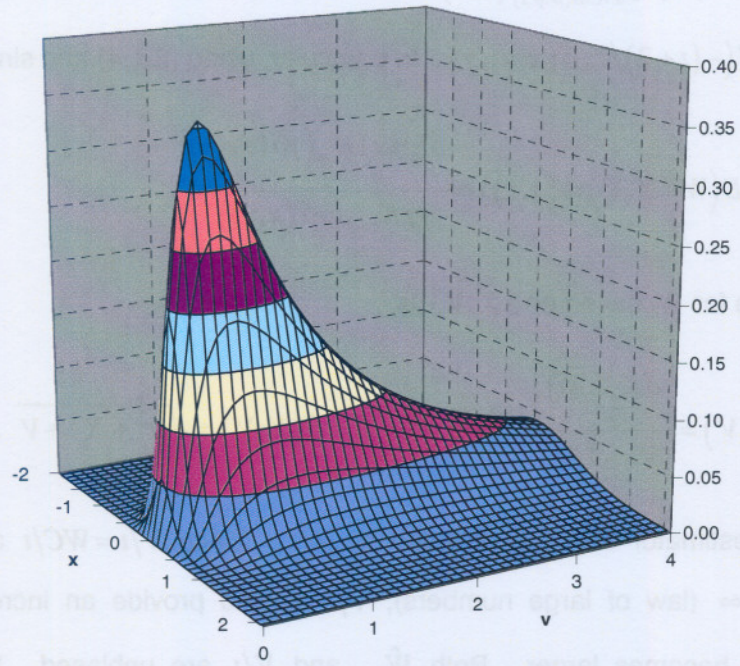

Figure 2.15: The $SNCIG(0,0.5,10)$-density.

$$
\begin{aligned}
&\frac{1}{\sqrt{w}}\varphi\left(\frac{x-\beta w}{\sqrt{w}}\right)\frac{1}{w}f_{\chi^2_\iota}\left(\frac{v}{w}\right)f_{UIG(\psi)}(w)\\
&=\frac{1}{\sqrt{w}\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\beta w}{\sqrt{w}}\right)^2\right)\frac{1}{w}\frac{\left(\frac{v}{w}\right)^{\frac{\iota}{2}-1}\exp\left(-\frac{v}{2w}\right)}{\Gamma\left(\frac{\iota}{2}\right)2^{\frac{\iota}{2}}}\frac{\psi e^{\psi^2}}{\sqrt{2\pi}}w^{-\frac{3}{2}}\exp\left(-\frac{\psi^2}{2}\left(w^{-1}+w\right)\right)\\
&=\frac{\psi\exp\left(\psi^2+\beta x\right)v^{\frac{\iota}{2}-1}}{\pi\Gamma\left(\frac{\iota}{2}\right)2^{\frac{\iota}{2}}}\left(\frac{\gamma}{\delta(x,v)}\right)^{\frac{\iota}{2}+1}\frac{K_{\frac{\iota}{2}+1}\left(\gamma\delta(x,v)\right)}{w^{\frac{\iota}{2}+2}}\left(\frac{\delta(x,v)}{\gamma}\right)^{\frac{\iota}{2}+1}\frac{\exp\left(-\frac{1}{2}\left(\frac{\delta(x,v)}{w}+\gamma w\right)\right)}{2K_{\frac{\iota}{2}+1}\left(\gamma\delta(x,v)\right)}\\
&=f_{SNCIG(\beta,\psi,\iota)}(x,v)\,f_{GIG\left(-\frac{\iota+2}{2},\delta(x,v),\gamma\right)}(w)
\end{aligned}
\tag{2.5.5}
$$

where

$$
\delta(x,v)=\sqrt{\psi^2+x^2+v}\ \text{ and }\ \gamma=\sqrt{\psi^2+\beta^2}\ . \tag{2.5.6}
$$

As noted in (2.5.2) the first (and thus also the last) line of (2.5.5) is the joint density of (X,V,W). Therefore, dividing (2.5.5) by $f_{SNCIG(\beta,\psi,\iota)}(x,v)$, we find that the conditional distribution of W given $(X,V)=(x,v)$ is $GIG\left(-(\iota+2)/2,\delta(x,v),\gamma\right)$. In particular, using (2.1.4) and since

$$
E\left(W\middle|(X,V)=(x,v)\right)=\frac{\delta(x,v)K_{\frac{\iota}{2}}\left(\gamma\delta(x,v)\right)}{\gamma K_{\frac{\iota+2}{2}}\left(\gamma\delta(x,v)\right)} \tag{2.5.7}
$$

the MSE best estimate for W based on (X,V) is

$$
\hat{W}_{XV}=E\left(W\middle|X,V\right)=\frac{\delta K_{\frac{\iota}{2}}(\gamma\delta)}{\gamma K_{\frac{\iota+2}{2}}(\gamma\delta)}\ \text{where now}\ \delta=\delta(X,V)=\sqrt{\psi^2+X^2+V}\ . \tag{2.5.8}
$$

A simple alternative estimator of W is provided by V/ι. Since $V/\iota=WC/\iota$ and $C/\iota\to 1$ with probability 1 as $\iota\to\infty$ (law of large numbers), V/ι should provide an increasingly accurate estimate of W as ι becomes larger. Both $\hat{W}_{XV}$ and V/ι are unbiased. We computed by

simulation the MSE ratio $E\left(\hat{W}_{XV}-W\right)^2 \Big/ Var(W)$ of $\hat{W}_{XV}$ and show it in Table 2.3. The MSE ratio of V/ι is given by

$$\frac{E\left(\frac{V}{\iota}-W\right)^2}{Var(W)} = \frac{E\left(\frac{WC}{\iota}-W\right)^2}{Var(W)} = \frac{E\left(W^2\right)\frac{1}{\iota^2}E(C-\iota)^2}{Var(W)}$$

$$= \frac{\left(1+\frac{1}{\psi^2}\right)\frac{2\iota}{\iota^2}}{1/\psi^2} = \frac{2}{\iota}\left(1+\psi^2\right) \qquad (2.5.9)$$

which does not depend on β and is given in the last column of Table 2.3. It is clear that V/ι improves with smaller values of ψ and larger values of ι. Indeed, when $\iota \to \infty$, V/ι converges to W. $\hat{W}_{XV}$ is a better estimator than V/ι, especially for small values of ι. It also improves with smaller values of ψ and larger values of ι and, in addition, it improves when β moves away from 0 (since dependence is symmetric in β, Table 2.3 is restricted to non-negative β's).

Analogous to the discussion in Section 2.3.5, just below (2.3.19), we will now inspect the lower boundary of $\hat{W}_{XV}$ given in (2.5.8). From (2.5.6) and $V \geq 0$ it follows that $\delta = \delta(X,V)$, δ/γ and $\gamma\delta$ are all monotone increasing in X and V. Plotting the ratio $K_{\iota/2}(z)\big/K_{(\iota+2)/2}(z)$ for different values of ι yields graphs similar to Figure 2.8, which means that $K_{\iota/2}(z)\big/K_{(\iota+2)/2}(z)$ is also monotone increasing in z. Thus, $\hat{W}_{XV}$ given in (2.5.8) reaches a minimum, which is the smallest possible estimate $\hat{W}_{XV}$ for W, in the argument $X = V = 0$, i.e. $\hat{W}_{XV} \geq \psi K_{\iota/2}(\gamma\psi)\big/\gamma K_{(\iota+2)/2}(\gamma\psi)$ since $\delta = \psi$.

Table 2.4 gives these minimum values for different values of β, ψ and ι. We see that the minimum values of $\hat{W}_{XV}$ become smaller for larger values of β and ι and smaller values of ψ. It is also clear that ψ and ι has a much greater impact on the minimum values than β. When we compare (2.3.18) with (2.5.8) we see that the minimum values of $\hat{W}_X$ become equal to the minimum values of $\hat{W}_{XV}$ in the limiting case when $\iota \to 0$. This can be seen when comparing the minimum values in Table 2.2 with their corresponding values in Table 2.4 as ι gets smaller. For instance, for $\psi = 2$ and $\beta = 0$, the minimum value of $\hat{W}_X$ (effectively $\iota = 0$) is 0.894 while the minimum values of $\hat{W}_{XV}$ are 0.717, 0.344 and 0.079 respectively for $\iota = 2$, $\iota = 10$ and $\iota = 50$.

Referring to Figure 2.1 we see that, as argued earlier, a $UIG(2)$-distributed variable (such as W in this case) frequently takes on values smaller than 0.894, so that we can expect $\hat{W}_X$ to estimate W rather poorly. However, as ι becomes larger, the lower boundary of $\hat{W}_{XV}$ becomes lower. Figure 2.1 shows that $W \sim UIG(2)$ would take on a value of less than 0.344 ($\iota = 10$) far less often than 0.717 ($\iota = 2$), and would hardly ever take on a value of less than 0.079 ($\iota = 50$). The same reasoning holds for the other values of β and ψ. Thus, provided ι is not too small, say $\iota \geq 20$, we would not expect this lower bound of $\hat{W}_{XV}$ to come into play and radically reduce the accuracy with which $\hat{W}_{XV}$ estimates W.

ψ	ι	$\hat{W}_{XV}$					$\frac{V}{\iota}$
		β=0	β=0.25	β=0.5	β=1	β=2	
0.10	2	0.348	0.092	0.029	0.009	0.002	1.0100
	10	0.132	0.058	0.024	0.007	0.002	0.2020
	50	0.036	0.024	0.014	0.006	0.002	0.0404
0.25	2	0.334	0.220	0.121	0.044	0.014	1.0625
	10	0.143	0.117	0.073	0.033	0.012	0.2125
	50	0.039	0.035	0.028	0.018	0.008	0.0425
0.50	2	0.393	0.345	0.257	0.135	0.049	1.2500
	10	0.167	0.155	0.130	0.084	0.038	0.2500
	50	0.046	0.044	0.041	0.033	0.021	0.0500
1.00	2	0.540	0.521	0.470	0.339	0.163	2.0000
	10	0.251	0.245	0.231	0.190	0.115	0.4000
	50	0.072	0.070	0.069	0.064	0.051	0.0800
2.00	2	0.764	0.755	0.730	0.639	0.429	5.0000
	10	0.468	0.463	0.453	0.414	0.315	1.0000
	50	0.161	0.161	0.160	0.155	0.137	0.2000
5.00	2	0.946	0.943	0.937	0.912	0.822	26.0000
	10	0.825	0.824	0.817	0.799	0.728	5.2000
	50	0.506	0.503	0.502	0.494	0.466	1.0400

Table 2.3: MSE ratios of the estimates $\hat{W}_{XV}$ and V/ι of W given (X,V).

ψ	ι	β=0	β=0.25	β=0.5	β=1	β=5
0.25	2	0.031	0.031	0.031	0.030	0.021
	10	0.006	0.006	0.006	0.006	0.006
	50	0.001	0.001	0.001	0.001	0.001
0.50	2	0.119	0.118	0.115	0.108	0.061
	10	0.025	0.025	0.025	0.025	0.023
	50	0.005	0.005	0.005	0.005	0.005
1.00	2	0.370	0.367	0.356	0.325	0.150
	10	0.099	0.099	0.098	0.098	0.081
	50	0.020	0.020	0.020	0.020	0.020
2.00	2	0.717	0.713	0.702	0.662	0.325
	10	0.344	0.344	0.342	0.335	0.231
	50	0.079	0.079	0.079	0.079	0.076
5.00	2	0.943	0.942	0.938	0.926	0.678
	10	0.807	0.806	0.804	0.795	0.607
	50	0.412	0.412	0.412	0.410	0.364

Table 2.4: Lower bounds for $\hat{W}_{XV}$ for different values of β, ψ and ι.

We may also want to estimate Z and C from (X,V). From the conditional distribution of W given (X,V) and (2.1.4) we can write the MSE estimator of $Z=(X-\beta W)/\sqrt{W}$ as

$$\begin{aligned}\hat{Z}_{XV} &= E(Z|X,V) = XE\left(W^{-\frac{1}{2}}|X,V\right) - \beta E\left(W^{\frac{1}{2}}|X,V\right) \\ &= X\left(\frac{\gamma}{\delta}\right)^{\frac{1}{2}} \frac{K_{\frac{\iota+3}{2}}(\gamma\delta)}{K_{\frac{\iota+2}{2}}(\gamma\delta)} - \beta\left(\frac{\delta}{\gamma}\right)^{\frac{1}{2}} \frac{K_{\frac{\iota+1}{2}}(\gamma\delta)}{K_{\frac{\iota+2}{2}}(\gamma\delta)}\end{aligned} \tag{2.5.10}$$

and the MSE estimator of $C=V/W$ as

$$\begin{aligned}\hat{C}_{XV} &= E(C|X,V) = VE\left(W^{-1}|X,V\right) \\ &= V\frac{\gamma}{\delta}\frac{K_{\frac{\iota+4}{2}}(\gamma\delta)}{K_{\frac{\iota+2}{2}}(\gamma\delta)}\end{aligned} \tag{2.5.11}$$

Tables 2.5 and 2.6 show the corresponding MSE ratios. Here $\hat{Z}_{XV}$ tends to estimate Z more accurately for smaller β, larger ψ and larger ι while $\hat{C}_{XV}$ tends to estimate C more accurately for larger β, larger ψ and smaller ι. Overall $\hat{C}_{XV}$ provides a much poorer estimate than $\hat{W}_{XV}$ and $\hat{Z}_{XV}$.

ψ	ι	β=0	β=0.25	β=0.5	β=1	β=2
0.10	2	0.115	0.129	0.148	0.183	0.249
	10	0.041	0.047	0.058	0.081	0.126
	50	0.010	0.012	0.016	0.026	0.049
0.25	2	0.112	0.131	0.169	0.245	0.376
	10	0.040	0.049	0.069	0.116	0.212
	50	0.010	0.012	0.018	0.036	0.084
0.50	2	0.101	0.121	0.169	0.288	0.478
	10	0.039	0.048	0.071	0.139	0.290
	50	0.010	0.012	0.018	0.042	0.111
1.00	2	0.074	0.090	0.135	0.266	0.510
	10	0.035	0.043	0.066	0.142	0.331
	50	0.009	0.012	0.018	0.043	0.125
2.00	2	0.038	0.046	0.072	0.161	0.390
	10	0.024	0.030	0.047	0.108	0.286
	50	0.008	0.010	0.017	0.040	0.121
5.00	2	0.009	0.011	0.018	0.043	0.134
	10	0.008	0.010	0.016	0.038	0.118
	50	0.005	0.006	0.010	0.024	0.077

Table 2.5: MSE ratios of the estimate $\hat{Z}_{XV}$ of Z given (X,V).

ψ	ι	β=0	β=0.25	β=0.5	β=1	β=2
0.10	2	0.661	0.654	0.639	0.607	0.552
	10	0.855	0.850	0.840	0.821	0.779
	50	0.962	0.962	0.957	0.946	0.925
0.25	2	0.646	0.633	0.606	0.548	0.449
	10	0.849	0.841	0.825	0.782	0.698
	50	0.961	0.959	0.953	0.936	0.890
0.50	2	0.593	0.582	0.550	0.477	0.354
	10	0.822	0.816	0.799	0.740	0.613
	50	0.955	0.952	0.947	0.923	0.858
1.00	2	0.457	0.450	0.432	0.377	0.266
	10	0.735	0.729	0.714	0.659	0.522
	50	0.927	0.926	0.919	0.896	0.821
2.00	2	0.257	0.256	0.250	0.229	0.174
	10	0.522	0.520	0.513	0.481	0.392
	50	0.833	0.832	0.829	0.809	0.742
5.00	2	0.068	0.068	0.068	0.066	0.060
	10	0.181	0.180	0.179	0.176	0.162
	50	0.492	0.494	0.491	0.483	0.458

Table 2.6: MSE ratios of the estimate $\hat{C}_{XV}$ of C given (X,V).

In order to use $\hat{W}_{XV}$, $\hat{Z}_{XV}$ and $\hat{C}_{XV}$ for model checking purposes we would need their distributions. $\hat{W}_{XV}$, $\hat{Z}_{XV}$ and $\hat{C}_{XV}$ would have been $UIG(\psi)$, unit normal and χ^2_t-distributed respectively if they were very accurate, but we cannot rely on this. Computing their exact theoretical distributions seems quite hard but we can, as in Section 2.3.5, estimate these distributions empirically as accurately as required by generating enough realisations of $\hat{W}_{XV}$, $\hat{Z}_{XV}$ and $\hat{C}_{XV}$ at given β, ψ and t. This generation process involves the following steps:

- Generate $Z \sim N(0,1)$, $C \sim \chi^2_t$ and $W \sim UIG(\psi)$ (see Section 2.8).
- Set $X = \beta W + \sqrt{W}Z$ and $V = WC$.
- Compute $\hat{W}_{XV}$, $\hat{Z}_{XV}$ and $\hat{C}_{XV}$ from (2.5.8), (2.5.10) and (2.5.11).
- Repeat as many times as required.

2.6 The TIG Distribution

2.6.1 Definition

Extensions to skewed forms of the t-distribution have been studied by various authors. Fernándes and Steel (1998) propose a general procedure for introducing skewness into symmetric distributions and apply this procedure to a t-distribution to generate a skew-t distribution. Azzalini and Capitanio (1999, 2003) show that a skew-t distribution can be obtained from a skew-normal distribution in the same way that a t-distribution can be obtained from a normal distribution. Jones and Faddy (2003) study the effects of skewing the t-distribution using the Beta distribution.

The definitions of the previous sections and that of the t-distribution suggest a new skewed t-distribution, which we will now introduce. We shall call the distribution of

$$T = \frac{X}{\sqrt{V/t}} = \frac{\beta\sqrt{W} + Z}{\sqrt{C/t}}, \tag{2.6.1}$$

with X, W, Z, C and V as above, the "t inverse Gaussian distribution" (abbreviated $TIG(\beta, \psi, t)$). Notice that if $\beta = 0$ then this distribution does not depend on the parameter ψ and becomes the ordinary t-distribution with t degrees of freedom. In general, conditionally given $W = w$, $\beta\sqrt{W} + Z$ is $N(\beta\sqrt{w}, 1)$-distributed so that T is non-centrally t-distributed with conditional density in the argument $T = t$ (Johnson, Kotz and Balakrishnan (1995), page 516)

$$f_{T|W=w}(t)=\frac{\exp\left(-\frac{1}{2}\beta^2 w\right)}{\sqrt{\pi\iota}\Gamma(\iota/2)}\left(\frac{\iota}{\iota+t^2}\right)^{\frac{\iota+1}{2}}\sum_{r=0}^{\infty}\frac{\Gamma\left(\frac{\iota+r+1}{2}\right)}{r!}\left(\frac{t\beta\sqrt{2w}}{\sqrt{\iota^2+t^2}}\right)^r$$

$$=\exp\left(-\frac{1}{2}\beta^2 w\right)\sum_{r=0}^{\infty}\frac{\beta^r}{r!}\frac{w^{\frac{r}{2}}2^{\frac{r}{2}}}{\iota^{\frac{r+2}{2}}}\frac{\Gamma\left(\frac{\iota+r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\iota}{2}\right)}\frac{t^r}{\left(1+\frac{t^2}{\iota}\right)^{\frac{\iota+r+1}{2}}} \tag{2.6.2}$$

Multiplying by $f_{UIG(\psi)}(w)$ and integrating out w, we obtain the unconditional density as

$$f_{TIG(\beta,\psi,\iota)}(t)$$

$$=\int_0^{\infty} f_{T|W=w}(t)\,f_{UIG(\psi)}(w)\,dw$$

$$=\sum_{r=0}^{\infty}\frac{\beta^r}{r!}\frac{2^{\frac{r}{2}}}{\iota^{\frac{r+2}{2}}}\frac{\Gamma\left(\frac{\iota+r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\iota}{2}\right)}\frac{t^r\psi\exp(\psi^2)}{\left(1+\frac{t^2}{\iota}\right)^{\frac{\iota+r+1}{2}}\sqrt{2\pi}}\int_0^{\infty}w^{\frac{r-3}{2}}\exp\left(-\frac{1}{2}\left(\frac{\psi^2}{w}+\left(\psi^2+\beta^2\right)w\right)\right)dw$$

$$=\frac{\psi\exp(\psi^2)}{\pi}\sum_{r=0}^{\infty}\frac{\beta^r}{r!}\frac{2^{\frac{r+1}{2}}}{\iota^{\frac{r+2}{2}}}\frac{\Gamma\left(\frac{\iota+r+1}{2}\right)}{\Gamma\left(\frac{\iota}{2}\right)}\frac{\left(\frac{\psi}{\sqrt{\psi^2+\beta^2}}\right)^{\frac{r-1}{2}}t^r}{\left(1+\frac{t^2}{\iota}\right)^{\frac{\iota+r+1}{2}}}K_{\frac{r-1}{2}}\left(\psi\sqrt{\psi^2+\beta^2}\right) \tag{2.6.3}$$

2.6.2 Properties

In order to evaluate the expectations of the $TIG(\beta,\psi,\iota)$ distribution, we need a few intermediate results. Firstly, the Gamma function (see Rice (1995), page 50) is defined as

$$\Gamma(a)=\int_0^{\infty}c^{a-1}\exp(-c)dc<\infty\Leftrightarrow a>0\ . \tag{2.6.4}$$

Secondly, note that from (2.4.1) we can write

$$E\left(C^{-k/2}\right)=\int_0^{\infty}\frac{c^{-k/2}c^{\iota/2-1}\exp(-c/2)}{\Gamma\left(\frac{\iota}{2}\right)2^{\frac{\iota}{2}}}dc=\frac{1}{\Gamma\left(\frac{\iota}{2}\right)2^{\frac{k}{2}+1}}\int_0^{\infty}\left(\frac{c}{2}\right)^{\frac{\iota-k}{2}-1}\exp\left(-\frac{c}{2}\right)dc\ . \tag{2.6.5}$$

Comparing (2.6.4) and (2.6.5) we see that $E\left(C^{-k/2}\right)$ will exist if and only if $\iota>k$. Therefore, from the definition of T and the facts that W and C are independent and $C>0$, we can write

$$E\left(|T|^k\right)=E\left(\left|\beta\sqrt{W}+Z\right|^k\left|\frac{C}{\iota}\right|^{-k/2}\right)=\iota^{\frac{k}{2}}E\left(\left|\beta\sqrt{W}+Z\right|^k\right)E\left(C^{-\frac{k}{2}}\right)<\infty\Leftrightarrow\iota>k\ , \tag{2.6.6}$$

which means that $E\left(T^k\right)$ will exist if and only if $\iota>k$. Provided these expectations exist, we can use (2.1.4), (1.6.3) and the fact that

$$\begin{aligned}E\left(C^{-\frac{1}{2}}\right)&=\int_0^\infty\frac{c^{\frac{\iota-3}{2}}\exp(-c/2)}{\Gamma\left(\frac{\iota}{2}\right)2^{\frac{\iota}{2}}}dc=\frac{1}{\Gamma\left(\frac{\iota}{2}\right)2^{\frac{3}{2}}}\int_0^\infty\left(\frac{c}{2}\right)^{\frac{\iota-3}{2}}\exp(-c/2)dc=\frac{\Gamma\left(\frac{\iota-1}{2}\right)}{\sqrt{2}\Gamma\left(\frac{\iota}{2}\right)}\\E\left(C^{-1}\right)&=\int_0^\infty\frac{c^{\frac{\iota-4}{2}}\exp(-c/2)}{\Gamma\left(\frac{\iota}{2}\right)2^{\frac{\iota}{2}}}dc=\frac{1}{4\Gamma\left(\frac{\iota}{2}\right)}\int_0^\infty\left(\frac{c}{2}\right)^{\frac{\iota-4}{2}}\exp(-c/2)dc=\frac{\Gamma\left(\frac{\iota}{2}-1\right)}{2\Gamma\left(\frac{\iota}{2}\right)}=\frac{1}{\iota-2}\end{aligned} \tag{2.6.7}$$

to write the first two moments of T as

$$\begin{aligned}E(T)&=E\left(\beta\sqrt{W}+Z\right)E\left((C/\iota)^{-1/2}\right)=\sqrt{\frac{\iota}{\pi}}\beta\psi\exp\left(\psi^2\right)K_0\left(\psi^2\right)\Gamma\left(\frac{\iota-1}{2}\right)\Big/\Gamma\left(\frac{\iota}{2}\right)\\E\left(T^2\right)&=E\left(\beta^2W+2\beta\sqrt{W}Z+Z^2\right)E\left((C/\iota)^{-1}\right)=\frac{\iota\left(\beta^2+1\right)}{\iota-2}\ .\end{aligned} \tag{2.6.8}$$

The variance of T is then given by

$$Var(T)=\frac{\iota\left(\beta^2+1\right)}{\iota-2}-\frac{\iota}{\pi}\beta^2\psi^2\exp\left(2\psi^2\right)K_0^2\left(\psi^2\right)\left(\Gamma\left(\frac{\iota-1}{2}\right)\Big/\Gamma\left(\frac{\iota}{2}\right)\right)^2\ . \tag{2.6.9}$$

Figures 2.16 and 2.17, with 2.18 and 2.19 on a log-scale, show $TIG(\beta,\psi,\iota)$-densities for different values of the parameters. From (2.6.3) it can easily be seen that $f_{TIG(-\beta,\psi,\iota)}(t)=f_{TIG(\beta,\psi,\iota)}(-t)$. This means that the density of $TIG(\beta,\psi,\iota)$ can be obtained from the density of $TIG(-\beta,\psi,\iota)$ by flipping the graph around the $t=0$ axis. Therefore, we show only graphs for $\beta>0$. Once again we see that β shifts the density along the t-axis, while smaller ψ-values and larger ι-values lead to densities that are more peaked.

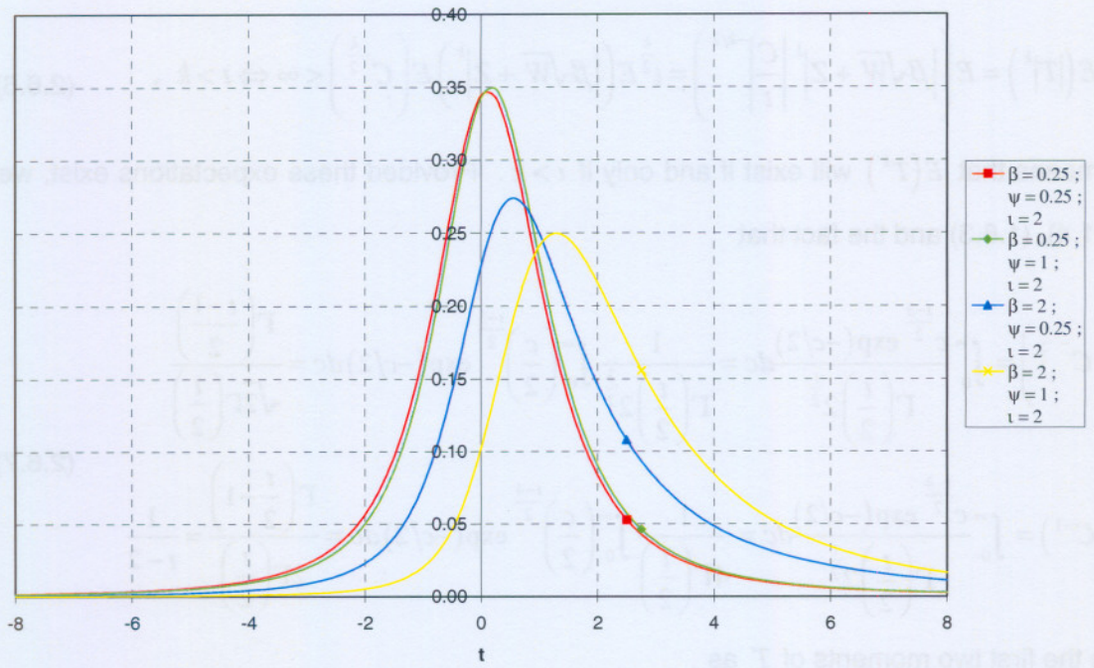

Figure 2.16: $TIG(\beta,\psi,2)$-densities.

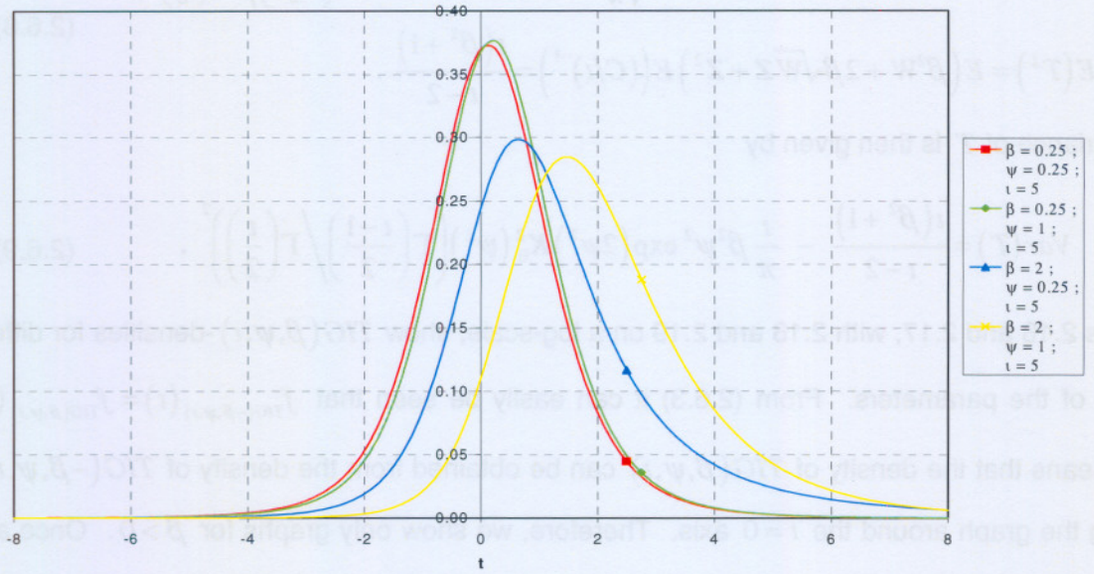

Figure 2.17: $TIG(\beta,\psi,5)$-densities.

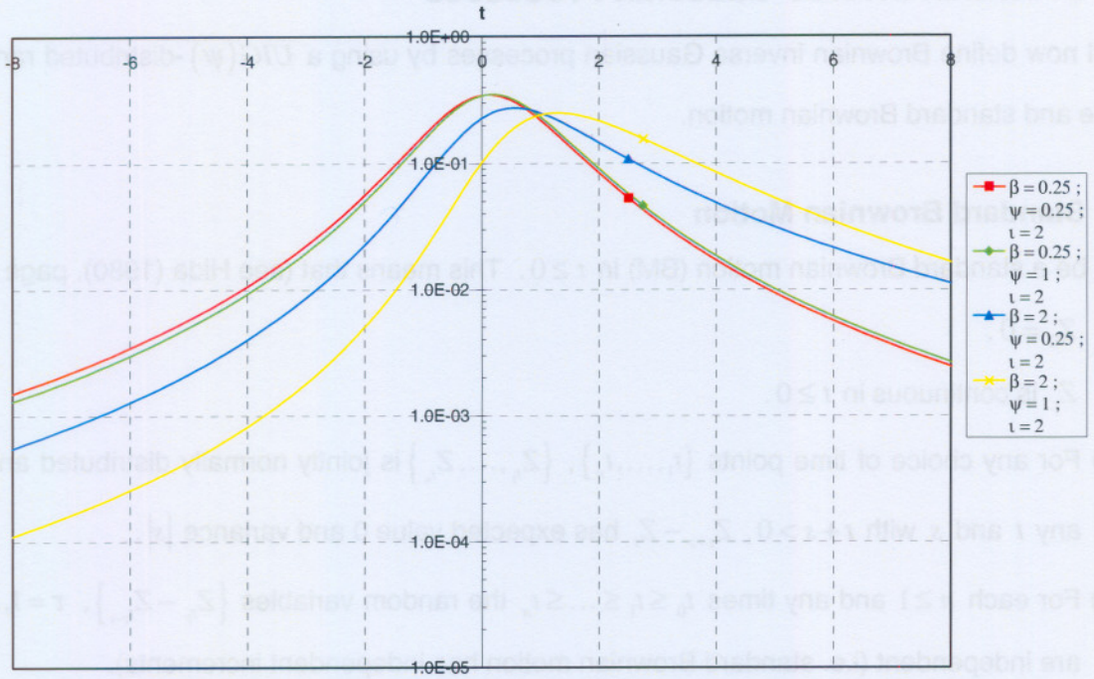

Figure 2.18: $TIG(\beta,\psi,2)$-densities on a log-scale.

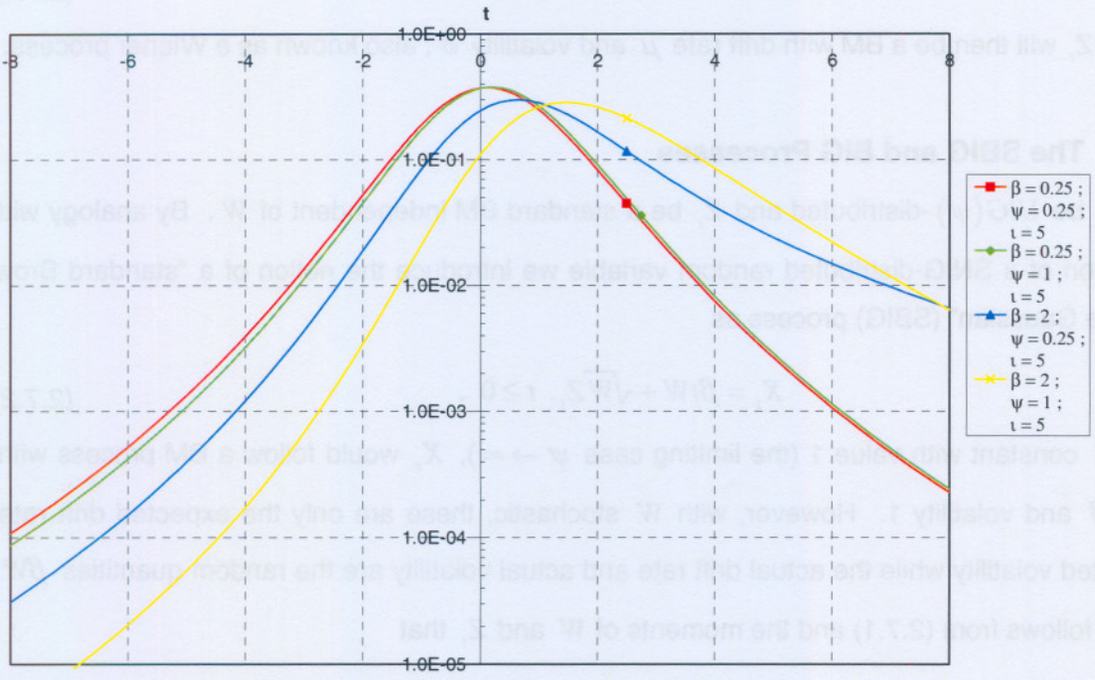

Figure 2.19: $TIG(\beta,\psi,5)$-densities on a log-scale.

2.7 Brownian Inverse Gaussian Processes

We will now define Brownian inverse Gaussian processes by using a $UIG(\psi)$-distributed random variable and standard Brownian motion.

2.7.1 Standard Brownian Motion

Let Z_t be a standard Brownian motion (BM) in $t \geq 0$. This means that (see Hida (1980), page 44)

(i) $Z_0 = 0$.

(ii) Z_t is continuous in $t \geq 0$.

(iii) For any choice of time points $\{t_1,\ldots,t_n\}$, $\{Z_{t_1},\ldots,Z_{t_n}\}$ is jointly normally distributed and for any t and s with $t+s>0$, $Z_{t+s}-Z_t$ has expected value 0 and variance $|s|$.

(iv) For each $n \geq 1$ and any times $t_0 \leq t_1 \leq \ldots \leq t_n$ the random variables $\{Z_{t_\tau} - Z_{t_{\tau-1}}\}$, $\tau = 1,\ldots,n$ are independent (i.e. standard Brownian motion has independent increments).

(v) For any $s,t>0$,

$$Cov(Z_t, Z_s) = \min(t,s) \ . \tag{2.7.1}$$

$\mu t + \sigma Z_t$ will then be a BM with drift rate μ and volatility σ, also known as a Wiener process.

2.7.2 The SBIG and BIG Processes

Let W be $UIG(\psi)$-distributed and Z_t be a standard BM independent of W. By analogy with the definition of a SNIG-distributed random variable we introduce the notion of a "standard Brownian inverse Gaussian" (SBIG) process as

$$X_t = \beta t W + \sqrt{W} Z_t, \ t \geq 0 \ . \tag{2.7.2}$$

For W constant with value 1 (the limiting case $\psi \to \infty$), X_t would follow a BM process with drift rate β and volatility 1. However, with W stochastic, these are only the expected drift rate and expected volatility while the actual drift rate and actual volatility are the random quantities βW and W. It follows from (2.7.1) and the moments of W and Z_t that

$$E(X_t) = \beta t E(W) + E\left(W^{1/2}\right) E(Z_t) = \beta t \tag{2.7.3}$$

and

$$
\begin{aligned}
&E(X_t X_s) \\
&= E\left(\left(\beta t W + W^{1/2} Z_t\right)\left(\beta s W + W^{1/2} Z_s\right)\right) \\
&= \beta^2 ts E\left(W^2\right) + \beta t E\left(W^{3/2}\right) E(Z_s) + \beta s E\left(W^{3/2}\right) E(Z_t) + E(W) E(Z_t Z_s) \\
&= \beta^2 ts\left(1 + \frac{1}{\psi^2}\right) + \min(t,s)
\end{aligned}
\tag{2.7.4}
$$

so that

$$
\begin{aligned}
Cov(X_t, X_s) &= E(X_t X_s) - E(X_t) E(X_s) \\
&= \beta^2 ts\left(1 + \frac{1}{\psi^2}\right) + \min(t,s) - \beta^2 ts \\
&= \frac{\beta^2 ts}{\psi^2} + \min(t,s) \ .
\end{aligned}
\tag{2.7.5}
$$

In the symmetric case $\beta = 0$ we see that the SBIG process has the same covariance kernel as the BM process. Adding a translation parameter μ (any real number) and scale parameter $\delta > 0$ to generalise the SBIG process, we introduce the notion of a Brownian inverse Gaussian (BIG) process as

$$
Y_t = \mu t + \delta X_t = (\mu + \delta\beta W) t + \delta\sqrt{W} Z_t, \ t \geq 0 \ . \tag{2.7.6}
$$

Thus, Y_t follows a BM process with drift rate and volatility equal to the random quantities $\mu + \delta\beta W$ and $\delta^2 W$, with expected drift rate and expected volatility equal to $\mu + \delta\beta$ and δ^2 respectively. Using (2.7.3) and (2.7.6) we see that

$$
E(Y_t) = \mu t + \delta E(X_t) = (\mu + \delta\beta) t \tag{2.7.7}
$$

and, using (2.7.5),

$$
\begin{aligned}
Cov(Y_t, Y_s) &= Cov(\mu t + \delta X_t, \mu s + \delta X_s) \\
&= \delta^2 Cov(X_t, X_s) \\
&= \delta^2\left(\frac{\beta^2 ts}{\psi^2} + \min(t,s)\right) \\
&= \frac{\delta^2\beta^2 ts}{\psi^2} + \delta^2 \min(t,s) \ .
\end{aligned}
\tag{2.7.8}
$$

Figure 2.20 displays random Brownian motions with drift rate $\mu + \delta\beta$ and volatility δ^2, with $\mu = 0$ and $\beta = \delta = 1$. We generated random W's from a $UIG(\psi)$ distribution with $\psi = 1$ and calculated

the BIG processes corresponding to the Brownian motions shown in Figure 2.20. Figure 2.21 shows these BIG processes. We can clearly see the effect of W on the structural part and the volatility. Values of $W>1$ magnify volatility and structural deviations while values of $W<1$ diminish these deviations. For instance, the movement in the Brownian motion BM8 is magnified in the corresponding BIG process since $W=1.64$ while the movement in the Brownian motion BM4 is diminished in the corresponding BIG process since $W=0.33$. When W is close to 1 the Brownian motion is not changed very much, for example BM5.

2.7.3 Joint Distributions at Given Times

In Section 3.2 we shall model the return process on a stock in terms of a BIG process Y_t by considering a duration of 1 trading day. We model the start of the day as $t=0$ and market closure time as $t=1$. We shall then need the joint distribution of the returns at regular times throughout the trading day. In anticipation of those needs, divide the unit interval $[0,1]$ into I equal subintervals, each of length $\Delta=1/I$, with endpoints at times $0=t_0<t_1<t_2<\ldots<t_I=1$ where $t_i=i\Delta=i/I$. We will then be interested in the returns at the boundaries of these intervals, i.e. in $Y_{t_1},Y_{t_2},\ldots,Y_1$. We can obtain the joint distribution of these returns by noting, from (2.7.1), (2.7.2) and $0=t_0<t_1<t_2<\ldots<t_I=1$, that this joint distribution, conditionally given $W=w$, will be multivariate normal. Multiplying with the density of W and integrating out w gives us the unconditional joint density.

However, it is technically easier to approach this joint distribution by considering the increments $Y_{t_i}-Y_{t_{i-1}}$ since we can then exploit the independent increments property of the Brownian motion. Substituting from (2.7.6), let

$$D_i=Y_{t_i}-Y_{t_{i-1}}=(\mu+\delta\beta W)\Delta+\delta\sqrt{W\Delta}\tilde{Z}_i \text{ with } \tilde{Z}_i=\left(Z_{t_i}-Z_{t_{i-1}}\right)\Big/\sqrt{\Delta},\ i=1,2,\ldots,I \qquad (2.7.9)$$

and $\boldsymbol{D}=[D_1,D_2,\ldots,D_I]^T$. It follows from (iii) and (iv) of Section 2.7.1 that each $\tilde{Z}_i$ is $N(0,1)$-distributed and, given $W=w$, each D_i will be normally distributed with mean $(\mu+\delta\beta w)\Delta$ and variance $\delta^2 w\Delta$, independent of the other D_i's.

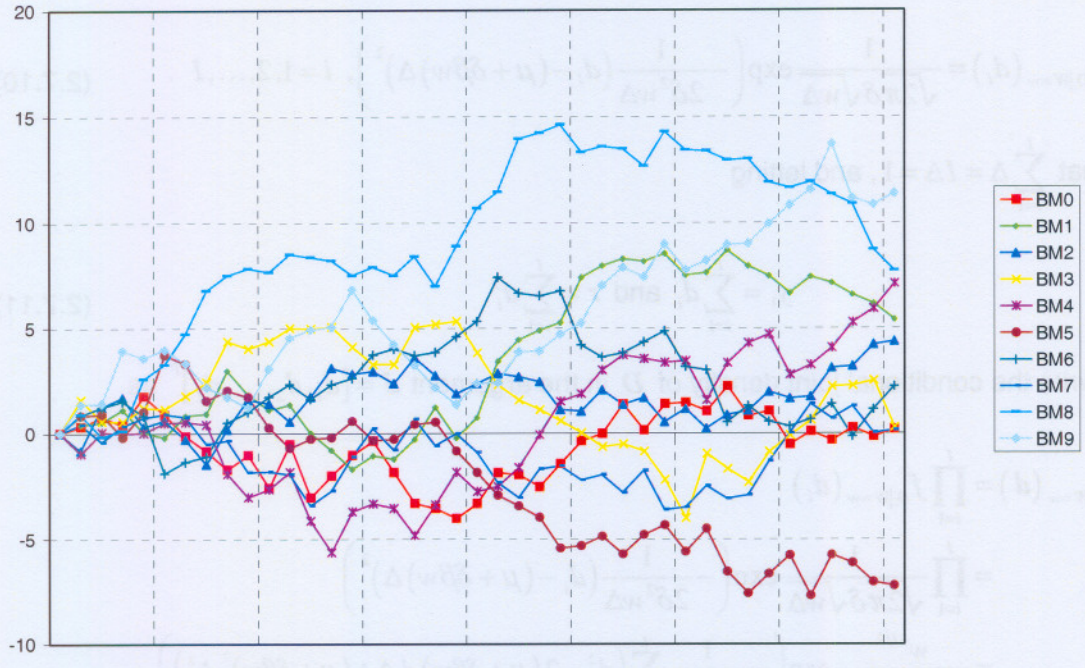

Figure 2.20: Random Brownian motions.

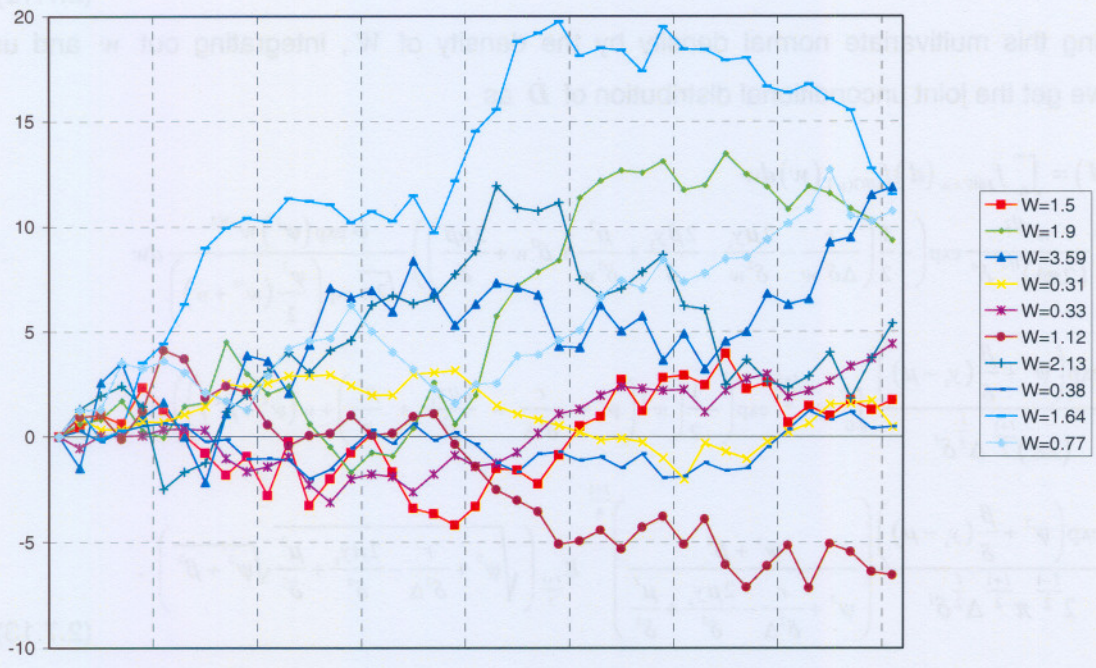

Figure 2.21: Corresponding BIG processes.

Thus, we can write the conditional density of D_i in the argument d_i as

$$f_{D_i|W=w}(d_i)=\frac{1}{\sqrt{2\pi}\delta\sqrt{w\Delta}}\exp\left(-\frac{1}{2\delta^2 w\Delta}\left(d_i-(\mu+\delta\beta w)\Delta\right)^2\right),\ i=1,2,\ldots,I\ . \tag{2.7.10}$$

Noting that $\sum_{i=1}^{I}\Delta=I\Delta=1$, and letting

$$y_1=\sum_{i=1}^{I}d_i \text{ and } r=\sum_{i=1}^{I}d_i^2 \tag{2.7.11}$$

we can write the conditional joint density of $\boldsymbol{D}$ in the argument $\boldsymbol{d}=[d_1,d_2,\ldots,d_I]^T$ as

$$\begin{aligned}
f_{\boldsymbol{D}|W=w}(\boldsymbol{d}) &= \prod_{i=1}^{I} f_{D_i|W=w}(d_i)\\
&= \prod_{i=1}^{I}\frac{1}{\sqrt{2\pi}\delta\sqrt{w\Delta}}\exp\left(-\frac{1}{2\delta^2 w\Delta}\left(d_i-(\mu+\delta\beta w)\Delta\right)^2\right)\\
&= \frac{w^{-I/2}}{(2\pi\Delta)^{I/2}\delta^I}\exp\left(-\frac{1}{2\delta^2 w\Delta}\sum_{i=1}^{I}\left(d_i^2-2(\mu+\delta\beta w)d_i\Delta+(\mu+\delta\beta w)^2\Delta^2\right)\right)\\
&= \frac{w^{-I/2}}{(2\pi\Delta)^{I/2}\delta^I}\exp\left(-\frac{1}{2\delta^2 w}\left(\frac{r}{\Delta}-2(\mu+\delta\beta w)y_1+(\mu+\delta\beta w)^2\right)\right).
\end{aligned} \tag{2.7.12}$$

Multiplying this multivariate normal density by the density of W, integrating out w and using (2.1.3) we get the joint unconditional distribution of $\boldsymbol{D}$ as

$$\begin{aligned}
f_{\boldsymbol{D}}(\boldsymbol{d}) &= \int_0^\infty f_{\boldsymbol{D}|W=w}(\boldsymbol{d})f_{UIG(\psi)}(w)dw\\
&= \int_0^\infty \frac{w^{-I/2}}{(2\pi\Delta)^{I/2}\delta^I}\exp\left(-\frac{1}{2}\left(\frac{r}{\Delta\delta^2 w}-\frac{2\mu y_1}{\delta^2 w}-\frac{2\beta y_1}{\delta}+\frac{\mu^2}{\delta^2 w}+\beta^2 w+\frac{2\mu\beta}{\delta}\right)\right)\frac{\psi\exp(\psi^2)w^{-3/2}}{\sqrt{2\pi}\exp\left(\frac{\psi^2}{2}(w^{-1}+w)\right)}dw\\
&= \frac{\psi\exp\left(\psi^2+\frac{\beta}{\delta}(y_1-\mu)\right)}{(2\pi)^{\frac{I+1}{2}}\Delta^{\frac{I}{2}}\delta^I}\int_0^\infty w^{-\frac{I+3}{2}}\exp\left(-\frac{1}{2}\left(w^{-1}\left(\psi^2+\frac{r}{\delta^2\Delta}-\frac{2\mu y_1}{\delta^2}+\frac{\mu^2}{\delta^2}\right)+w(\psi^2+\beta^2)\right)\right)dw\\
&= \frac{\psi\exp\left(\psi^2+\frac{\beta}{\delta}(y_1-\mu)\right)}{2^{\frac{I-1}{2}}\pi^{\frac{I+1}{2}}\Delta^{\frac{I}{2}}\delta^I}\left(\frac{\psi^2+\beta^2}{\psi^2+\frac{r}{\delta^2\Delta}-\frac{2\mu y_1}{\delta^2}+\frac{\mu^2}{\delta^2}}\right)^{\frac{I+1}{4}}K_{\frac{I+1}{2}}\left(\sqrt{\psi^2+\frac{r}{\delta^2\Delta}-\frac{2\mu y_1}{\delta^2}+\frac{\mu^2}{\delta^2}}\sqrt{\psi^2+\beta^2}\right).
\end{aligned} \tag{2.7.13}$$

We see that (2.7.13) depends upon the arguments d_i, $i=1,\ldots I$ only through their sum y_1 and sum of squares r, which means that the corresponding functions Y_1 and R constitute sufficient statistics for D_i, $i=1,\ldots,I$ in this context. Therefore, if we wish to base statistical inference on

D_i, $i=1,\ldots,I$, then we might as well base such inference on the sufficient statistics Y_1 and R. For this purpose we shall need their joint distribution. It is technically more convenient to work with $\tilde{V}=IS$ where S is the sum of squared deviations (SSD) given by

$$S=\sum_{i=1}^{I}\left(D_i-\frac{Y_1}{I}\right)^2=\sum_{i=1}^{I}D_i^2-\frac{2Y_1}{I}\sum_{i=1}^{I}D_i+\frac{Y_1^2}{I}=R-Y_1^2/I \tag{2.7.14}$$

(using (2.7.11)) rather than R. To calculate the joint density of Y_1 and $\tilde{V}$, let

$$Z=\sqrt{I}\bar{Z} \text{ with } \bar{Z}=\sum_{i=1}^{I}\tilde{Z}_i/I \tag{2.7.15}$$

and

$$\begin{aligned}C&=\sum_{i=1}^{I}\left(\tilde{Z}_i-\bar{Z}\right)^2\\&=\sum_{i=1}^{I}\tilde{Z}_i^2-2\bar{Z}\sum_{i=1}^{I}\tilde{Z}_i+I\bar{Z}^2\\&=\sum_{i=1}^{I}\tilde{Z}_i^2-I\bar{Z}^2\\&=\sum_{i=1}^{I}\tilde{Z}_i^2-Z^2\ .\end{aligned} \tag{2.7.16}$$

Since each $\tilde{Z}_i$ (see (2.7.9)) is $N(0,1)$-distributed, $\bar{Z}$ will be $N(0,1/I)$-distributed and Z will be $N(0,1)$-distributed (see Rice (1995), page 147), while C is χ^2_{I-1}-distributed independent of Z.

It follows from (2.7.9), (2.7.11) and (2.7.15) that

$$\begin{aligned}Y_1&=\mu\sum_{i=1}^{I}\Delta+\delta\beta W\sum_{i=1}^{I}\Delta+\delta\sqrt{W\Delta}\sum_{i=1}^{I}\tilde{Z}_i\\&=\mu+\delta\beta W+\delta\sqrt{W}Z\end{aligned} \tag{2.7.17}$$

so that (from (2.7.6))

$$X_1=\frac{Y_1-\mu}{\delta}=\beta W+\sqrt{W}Z\ . \tag{2.7.18}$$

From (2.7.9), (2.7.14) and (2.7.17) we see that

$$\tilde{V} = IS = I\sum_{i=1}^{I}\left(D_i - \frac{Y_1}{I}\right)^2 = \frac{1}{I}\sum_{i=1}^{I}\left(ID_i - Y_1\right)^2$$
$$= \frac{1}{I}\sum_{i=1}^{I}\left(\mu + \delta\beta W + \sqrt{I}\delta\sqrt{W}\tilde{Z}_i - \mu - \delta\beta W - \delta\sqrt{W}Z\right)^2$$
$$= \delta^2 W\sum_{i=1}^{I}\left(\tilde{Z}_i - Z/\sqrt{I}\right)^2$$
$$= \delta^2 W\sum_{i=1}^{I}\left(\tilde{Z}_i - \bar{Z}\right)^2$$
$$= \delta^2 WC\ . \qquad (2.7.19)$$

Finally, let

$$V = \tilde{V}/\delta^2 = WC\ . \qquad (2.7.20)$$

Comparing (2.7.18) and (2.7.20) to the definitions in Sections 2.3.1, 2.4.1 and 2.5.1, we see that the distribution of X_1 is $SNIG(\beta,\psi)$, the distribution of V is $SCIG(\psi, I-1)$ and the joint distribution of X_1 and V is $SNCIG(\beta,\psi,I-1)$. Therefore, using (1.5.5), the joint density of Y_1 and $\tilde{V}$ is given by

$$f_{Y_1,\tilde{V}}(y_1,\tilde{v}) = \frac{1}{\delta^3} f_{SNCIG(\beta,\psi,I-1)}\left(\frac{y_1-\mu}{\delta}, \frac{\tilde{v}}{\delta^2}\right)\ . \qquad (2.7.21)$$

2.7.4 Joint Distributions of BIG Maximum, Minimum and Values at Given Times

In Section 3.2.4 we shall be interested in using BIG processes to model returns of stocks. Historically, returns are often summarised in terms of open, high, low and close prices. For purposes arising from this situation we shall need to study the joint distribution of $Y = Y_1$ (see previous section),

$$A = \max_{0\le t\le 1}(Y_t) = \max_{0\le t\le 1}\left((\mu+\delta\beta W)t + \delta\sqrt{W}Z_t\right) \qquad (2.7.22)$$

and

$$B = -\min_{0\le t\le 1}(Y_t) = -\min_{0\le t\le 1}\left((\mu+\delta\beta W)t + \delta\sqrt{W}Z_t\right) \qquad (2.7.23)$$

for given parameters μ and δ. To calculate the joint density of (Y,A,B), let $\tilde{Y}_t$ denote a Brownian motion with drift rate η and volatility σ, let $\tilde{A} = \max_{0\le t\le 1}(\tilde{Y}_t)$, $\tilde{B} = -\min_{0\le t\le 1}(\tilde{Y}_t)$ and let $f_{\tilde{Y}_1,\tilde{A},\tilde{B}}\left(y,a,b\middle|\eta,\sigma^2\right)$ denote the joint density of $(\tilde{Y}_1,\tilde{A},\tilde{B})$ in the argument (y,a,b). Then the

conditional density of (Y,A,B) given $W=w$ is given by $f_{\tilde{Y}_1,\tilde{A},\tilde{B}}\left(y,a,b\middle|\mu+\delta\beta w,\delta^2 w\right)$ so that the conditional density of (Y,A,B) may be expressed as

$$f_{Y,A,B}\left(y,a,b\middle|\mu,\delta^2\right)=\int_0^\infty f_{\tilde{Y}_1,\tilde{A},\tilde{B}}\left(y,a,b\middle|\mu+\delta\beta w,\delta^2 w\right)f_{UIG(\psi)}(w)\,dw\ . \tag{2.7.24}$$

In order to calculate this integral we need an expression for $f_{\tilde{Y}_1,\tilde{A},\tilde{B}}\left(y,a,b\middle|\eta,\sigma^2\right)$. One possible expression is obtained by differentiating equation (81) on page 222 of Cox and Miller (1965) with respect to a and b. We have greater success in proceeding from the alternative equation (78) on page 222 of Cox and Miller (1965). The transition density for U_t with absorbing barriers at a and $-b$ is derived as

$$p(y_t,a,b,t)=\frac{1}{\sigma\sqrt{2\pi t}}\sum_{m=-\infty}^{\infty}\left(\exp\left(\frac{\eta y'_m}{\sigma^2}-\frac{(y_t-y'_m-\eta t)^2}{2\sigma^2 t}\right)-\exp\left(\frac{\eta y''_m}{\sigma^2}-\frac{(y_t-y''_m-\eta t)^2}{2\sigma^2 t}\right)\right) \tag{2.7.25}$$

where $y'_m=2m(a+b)$ and $y''_m=2a-y'_m$. Normalising (2.7.25) by setting $t=1$ and dropping it as an argument, denoting y_t by y and setting

$$\begin{aligned} f'_m&=\frac{\eta y'_m}{\sigma^2}-\frac{(y-y'_m-\eta)^2}{2\sigma^2}\\ f''_m&=\frac{\eta y''_m}{\sigma^2}-\frac{(y-y''_m-\eta)^2}{2\sigma^2}\end{aligned} \tag{2.7.26}$$

lead to

$$p(y,a,b)=\frac{1}{\sigma\sqrt{2\pi}}\sum_{m=-\infty}^{\infty}\left(\exp(f'_m)-\exp(f'')\right)\ . \tag{2.7.27}$$

Note that, in our current context,

$$p(y,a,b)=\frac{\partial}{\partial y}F_{\tilde{Y}_1,\tilde{A},\tilde{B}}\left(y,a,b\middle|\eta,\sigma^2\right) \tag{2.7.28}$$

where $F_{\tilde{Y}_1,\tilde{A},\tilde{B}}\left(y,a,b\middle|\eta,\sigma^2\right)=P\left(\tilde{Y}_1\le y,\tilde{A}\le a,\tilde{B}\le b\right)$, so that we still need to differentiate (2.7.27) with respect to a and b in order to obtain $f_{\tilde{Y}_1,\tilde{A},\tilde{B}}\left(y,a,b\middle|\eta,\sigma^2\right)$. Carrying out the differentiation we obtain

$$\frac{\partial p(y,a,b)}{\partial a} = \frac{1}{\sigma\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} \left(\exp(f_m') \frac{\partial f_m'}{\partial a} - \exp(f_m'') \frac{\partial f_m''}{\partial a} \right)$$
$$\frac{\partial^2 p(y,a,b)}{\partial b \partial a} = \frac{1}{\sigma\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} \left(\exp(f_m') \left(\frac{\partial f_m'}{\partial b} \frac{\partial f_m'}{\partial a} + \frac{\partial^2 f_m'}{\partial b \partial a} \right) - \exp(f_m') \left(\frac{\partial f_m''}{\partial b} \frac{\partial f_m''}{\partial a} + \frac{\partial^2 f_m''}{\partial b \partial a} \right) \right). \tag{2.7.29}$$

Substituting the derivatives

$$\begin{aligned}
\frac{\partial f_m'}{\partial a} &= \frac{\eta}{\sigma^2} 2m + \frac{(y - y_m' - \eta)}{\sigma^2} 2m = \frac{y - y_m'}{\sigma^2} 2m = \frac{\partial f_m'}{\partial b} \\
\frac{\partial^2 f_m'}{\partial b \partial a} &= \frac{2m}{\sigma^2}(-2m) = -\frac{4m^2}{\sigma^2} \\
\frac{\partial f_m''}{\partial a} &= \frac{\eta}{\sigma^2}(2-2m) + \frac{(y - y_m'' - \eta)}{\sigma^2}(2-2m) = \frac{y - y_m''}{\sigma^2}(2-2m) \\
\frac{\partial f_m''}{\partial b} &= \frac{\eta}{\sigma^2}(-2m) + \frac{(y - y_m'' - \eta)}{\sigma^2}(-2m) = -\frac{y - y_m''}{\sigma^2} 2m \\
\frac{\partial^2 f_m''}{\partial b \partial a} &= -\frac{(2-2m)}{\sigma^2}(-2m) = -\frac{4m(m-1)}{\sigma^2}
\end{aligned} \tag{2.7.30}$$

and noting from (2.7.26) that

$$\begin{aligned}
f_m' &= -\frac{1}{2\sigma^2}\left((y-\eta)^2 - 2y_m' y + y_m'^2 \right) \\
f_m'' &= -\frac{1}{2\sigma^2}\left((y-\eta)^2 - 2y_m'' y + y_m''^2 \right)
\end{aligned} \tag{2.7.31}$$

we see that

$$\begin{aligned}
& f_{\tilde{Y}_1, \tilde{A}, \tilde{B}}\left(y,a,b \middle| \eta, \sigma^2 \right) = \frac{\partial^2 p(y,a,b)}{\partial b \partial a} \\
&= \frac{1}{\sigma\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} \Bigg[\exp\left(-\frac{(y-\eta)^2 - 2y_m' y + y_m'^2}{2\sigma^2} \right) \left(\frac{4m^2 (y - y_m')^2}{\sigma^4} - \frac{4m^2}{\sigma^2} \right) - \\
&\qquad \exp\left(-\frac{(y-\eta)^2 - 2y_m'' y + y_m''^2}{2\sigma^2} \right) \left(\frac{4m(m-1)(y - y_m'')^2}{\sigma^4} - \frac{4m(m-1)}{\sigma^2} \right) \Bigg] \\
&= \frac{4\exp\left(-\frac{(y-\eta)^2}{2\sigma^2} \right)}{\sigma^5\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} \Bigg[\exp\left(-\frac{y_m'^2 - 2y_m' y}{2\sigma^2} \right) m^2 \left((y - y_m')^2 - \sigma^2 \right) - \\
&\qquad \exp\left(-\frac{y_m''^2 - 2y_m'' y}{2\sigma^2} \right) m(m-1) \left((y - y_m'')^2 - \sigma^2 \right) \Bigg].
\end{aligned} \tag{2.7.32}$$

Substituting (2.7.32) and (2.2.2) into (2.7.24), setting

$$\beta_\psi = \left(\beta^2+\psi^2\right)^{1/2}$$
$$\psi'_m = \left(\frac{(y-\mu)^2 + y'^2_m - 2y'_m y}{\delta^2} + \psi^2\right)^{1/2} \qquad (2.7.33)$$
$$\psi''_m = \left(\frac{(y-\mu)^2 + y''^2_m - 2y''_m y}{\delta^2} + \psi^2\right)^{1/2}$$

and carrying out the integration over w we obtain

$$\begin{aligned}
&f_{Y,A,B}\left(y,a,b\middle|\mu,\delta^2\right) \\
&= \int_0^\infty \frac{4}{\delta^5 w^{\frac{5}{2}}\sqrt{2\pi}} \sum_{m=-\infty}^{\infty}\left[\exp\left(-\frac{(y-\mu-\delta\beta w)^2}{2\delta^2 w}\right)\left\{\exp\left(-\frac{y'^2_m - 2y'_m y}{2\delta^2 w}\right)m^2\left((y-y'_m)^2-\delta^2 w\right)-\right.\right. \\
&\left.\left.\exp\left(-\frac{y''^2_m - 2y''_m y}{2\delta^2 w}\right)m(m-1)\left((y-y''_m)^2-\delta^2 w\right)\right\}\right]\frac{\psi\exp\left(\psi^2\right)}{\sqrt{2\pi}} w^{-\frac{3}{2}}\exp\left(-\frac{\psi^2}{2}\left(w^{-1}+w\right)\right)dw \\
&= \frac{2\psi\exp\left(\psi^2+\frac{\beta(y-\mu)}{\delta}\right)}{\pi\delta^5}\sum_{m=-\infty}^{\infty}\int_0^\infty\left[\exp\left(-\frac{1}{2}\left(\psi'_m w^{-1}+\beta_\psi^2 w\right)\right)m^2\left(\frac{(y-y'_m)^2}{w^4}-\frac{\delta^2}{w^3}\right)-\right. \\
&\qquad\left.\exp\left(-\frac{1}{2}\left(\psi''_m w^{-1}+\beta_\psi^2 w\right)\right)m(m-1)\left(\frac{(y-y''_m)^2}{w^4}-\frac{\delta^2}{w^3}\right)\right]dw \\
&= \frac{2\psi}{\pi\delta^5}\exp\left(\psi^2+\frac{\beta(y-\mu)}{\delta}\right)\sum_{m=-\infty}^{\infty}\left[m^2\left((y-y'_m)^2\left(\frac{\beta_\psi}{\psi'_m}\right)^3 K_3\left(\beta_\psi\psi'_m\right)-\delta^2\left(\frac{\beta_\psi}{\psi'_m}\right)^2 K_2\left(\beta_\psi\psi'_m\right)\right)\right. \\
&\qquad\left.-m(m-1)\left((y-y''_m)^2\left(\frac{\beta_\psi}{\psi''_m}\right)^3 K_3\left(\beta_\psi\psi''_m\right)-\delta^2\left(\frac{\beta_\psi}{\psi''_m}\right)^2 K_2\left(\beta_\psi\psi''_m\right)\right)\right].
\end{aligned} \qquad (2.7.34)$$

We will now show that $y'^2_m - 2y'_m y \geq 0$ and $y''^2_m - 2y''_m y \geq 0$ for all cases which ensures that the square roots in (2.7.33) exist, $\psi'_m > 0$, $\psi''_m > 0$ and makes the integration leading to (2.7.34) valid. From the definition of y, $a \geq 0$ and $b \geq 0$ it follows that $-b \leq y \leq a$ so that $-(a+b) \leq y \leq a+b$ and therefore $-m^2(a+b) \leq my \leq m^2 y \leq m^2(a+b)$. This means that $m^2(a+b) - my \geq 0$ for all $m \in \mathbb{Z}$, so that it follows from the definition of y'_m just below (2.7.25) that

$$y'^2_m - 2y'_m y = 4m^2(a+b)^2 - 4my(a+b) = 4(a+b)\left(m^2(a+b)-my\right) \geq 0\ . \qquad (2.7.35)$$

Furthermore, note that, for $m>0$, $a-m(a+b)\le 0$ and $a-m(a+b)-y=a-ma-mb-y\le 0$ since $-mb\le -b\le y$ while, for $m\le 0$, $a-m(a+b)\ge 0$ and $a-m(a+b)-y\ge 0$ since $y\le a$. This means that, for all $m\in\mathbb{Z}$, we see from the definition of y_m'' that

$$y_m''^2-2y_m''y=4\big(a-m(a+b)\big)\big(a-m(a+b)-y\big)\ge 0\ . \tag{2.7.36}$$

Note that in (2.7.34) the term $m=0$ does not contribute to the sum while only the first part contributes when $m=1$. We compute (2.7.34) by summing separately over the positive and negative values of m until convergence in each direction is achieved. Typically, this requires less than 10 terms in each range.

This concludes our derivation and discussion of the joint density of (Y,A,B). As mentioned above, the results will be applied in Section 3.2.4.

We next develop expressions for the joint densities of (Y,A) and of (Y,B) separately. Cox and Miller (1965) derive the transition density for $\tilde{Y}_t$ with absorbing barrier at a as equation (71) on page 221. Normalising by setting $t=1$, and dropping t as an argument, this is given by

$$p(y,a)=\frac{1}{\sigma\sqrt{2\pi}}\left(\exp\left(-\frac{(y-\eta)^2}{2\sigma^2}\right)-\exp\left(\frac{2\eta a}{\sigma^2}-\frac{(y-2a-\eta)^2}{2\sigma^2}\right)\right)\ . \tag{2.7.37}$$

Again, we still need to differentiate (2.7.37) with respect to a, leading to

$$\begin{aligned}
f_{\tilde{Y}_1,\tilde{A}}\left(y,a\middle|\eta,\sigma^2\right)&=\frac{\partial p(y,a)}{\partial a}\\
&=\frac{-1}{\sigma\sqrt{2\pi}}\exp\left(\frac{2\eta a}{\sigma^2}-\frac{(y-2a-\eta)^2}{2\sigma^2}\right)\left(\frac{2\eta}{\sigma^2}+\frac{2(y-2a-\eta)}{\sigma^2}\right)\\
&=\frac{\sqrt{2}(2a-y)}{\sigma^3\sqrt{\pi}}\exp\left(-\frac{1}{2\sigma^2}\Big((y-2a)^2-\eta(2y-\eta)\Big)\right)\ ,
\end{aligned} \tag{2.7.38}$$

so that

$$\begin{aligned}
&f_{Y,A}\left(y,a\middle|\mu,\delta^2\right)\\
&=\int_0^\infty f_{\tilde{Y}_1,\tilde{A}}\left(y,a\middle|\mu+\delta\beta w,\delta^2 w\right)f_{UIG(\psi)}(w)dw\\
&=\int_0^\infty\frac{\sqrt{2}(2a-y)}{\delta^3 w^{3/2}\sqrt{\pi}}\exp\left(-\frac{\big((y-2a)^2-(\mu-\delta\beta w)(2y-\mu-\delta\beta w)\big)}{2\delta^2 w}\right)\frac{\psi\exp(\psi^2)}{\sqrt{2\pi}w^{3/2}}\exp\left(-\frac{\psi^2(w^{-1}+w)}{2}\right)dw
\end{aligned}$$

$$= \frac{(2a-y)\psi \exp\left(\psi^2 + \frac{2y\beta}{\delta}\right)}{\pi\delta^3} \int_0^\infty w^{-3} \exp\left(-\frac{1}{2}\left(\left(\psi^2 + \frac{(y-\mu)^2 + 4a(a-y)}{\delta^2}\right)w^{-1} + \left(\psi^2 + \beta^2\right)w\right)\right) dw$$

$$= \frac{(2a-y)\psi \exp\left(\psi^2 + \frac{2y\beta}{\delta}\right)}{\pi\delta^3} \left(\frac{\beta_\psi}{\psi_a}\right)^2 K_2\left(\psi_a \beta_\psi\right) \qquad (2.7.39)$$

where $\psi_a = \left(\psi^2 + \left((y-\mu)^2 + 4a(a-y)\right)/\delta^2\right)^{1/2}$. Since $y \le a$ and $a \ge 0$, the square root exists and $\psi_a > 0$.

By symmetry, the joint density of (Y, B) can also be derived as above. We see that the joint densities of (Y, A) and of (Y, B) are much simpler to deal with than the full joint density of (Y, A, B); in particular, an infinite series such as that in (2.7.34) is not required. When modelling stock returns however, the low prices are usually also available if the high prices are available, and vice versa. Hence, the most relevant case from a practical point of view is the case based on all the return data (Y, A, B).

2.7.5 Estimating W from observations on extremes of X_t

Analogous to Sections 2.3.5 and 2.5.3, we shall find ourselves in the situation in Section 3.2 where we have observed extreme values of X_t, $0 \le t \le 1$, which can be modelled as having the distribution of $\beta t W + \sqrt{W} Z_t$ (see Section 2.7.2), and want to estimate W from these. If we use only X_1, we can compute $\hat{W}$ (and $\hat{Z} \equiv \hat{Z}_1$) as in Section 2.3.5 and use them as before. However, these estimates do not cater for the high and low data and in principle we expect to do better if we use that data as well. To achieve this, define $A_X = \max_{0 \le t \le 1}(X_t)$ and $B_X = -\min_{0 \le t \le 1}(X_t)$. In this case the minimum MSE estimator of W is $\hat{W}_{XAB} = E\left(W \middle| X_1, A_X, B_X\right)$ and we need the conditional distribution of W given $(X_1, A_X, B_X) = (x, a, b)$ to calculate this. In the context of (2.7.32), the variables (X_1, A_X, B_X, W) would have joint density $f_{\tilde{Y}_1, \tilde{A}, \tilde{B}}\left(x, a, b \middle| \beta w, w\right) f_{UIG(\psi)}(w)$. Thus, using (2.7.24) and a derivation very similar to (2.7.34) except for the power of w (and thus the indices of the Bessel functions), and setting $x'_m = 2m(a+b)$, $x''_m = 2a - x'_m$, $\beta_\psi = \left(\beta^2 + \psi^2\right)^{1/2}$, $\psi'_m = \left(x^2 + x'^2_m - 2x'_m x + \psi^2\right)^{1/2}$ and $\psi''_m = \left(x^2 + x''^2_m - 2x''_m x + \psi^2\right)^{1/2}$, we get

$$\hat{W}_{XAB} = E\left(W \middle| X_1, A_X, B_X\right)$$

$$= \int_0^\infty w f_{W|X_1=x,A_X=a,B_X=b}(w)dw$$

$$= \frac{\int_0^\infty w f_{\tilde{Y}_1,\tilde{A},\tilde{B}}\left(x,a,b\middle|\beta w, w\right) f_{UIG(\psi)}(w)dw}{\int_0^\infty f_{\tilde{Y}_1,\tilde{A},\tilde{B}}\left(x,a,b\middle|\beta w, w\right) f_{UIG(\psi)}(w)dw}$$

$$= \frac{\sum_{m=-\infty}^{\infty}\left(m^2 K'_{2,1} - m(m-1)K''_{2,1}\right)}{\sum_{m=-\infty}^{\infty}\left(m{K'_{3,2}}^2 - m(m-1)K''_{3,2}\right)}, \text{ where}$$

$$K'_{2,1} = (x-x'_m)^2\left(\frac{\beta_\psi}{\psi'_m}\right)^2 K_2\left(\beta_\psi\psi'_m\right) - \left(\frac{\beta_\psi}{\psi'_m}\right)K_1\left(\beta_\psi\psi'_m\right),$$

$$K''_{2,1} = (x-x''_m)^2\left(\frac{\beta_\psi}{\psi''_m}\right)^2 K_2\left(\beta_\psi\psi''_m\right) - \left(\frac{\beta_\psi}{\psi''_m}\right)K_1\left(\beta_\psi\psi''_m\right),$$

$$K'_{3,2} = (x-x'_m)^2\left(\frac{\beta_\psi}{\psi'_m}\right)^3 K_3\left(\beta_\psi\psi'_m\right) - \left(\frac{\beta_\psi}{\psi'_m}\right)^2 K_2\left(\beta_\psi\psi'_m\right) \text{ and} \quad (2.7.40)$$

$$K''_{3,2} = (x-x''_m)^2\left(\frac{\beta_\psi}{\psi''_m}\right)^3 K_3\left(\beta_\psi\psi''_m\right) - \left(\frac{\beta_\psi}{\psi''_m}\right)^2 K_2\left(\beta_\psi\psi''_m\right).$$

ψ	Steps	β=0	β=0.25	β=0.5	β=1	β=2
0.10	1000	0.178	0.084	0.020	0.009	0.005
	2000	0.170	0.071	0.025	0.007	0.005
	5000	0.173	0.067	0.026	0.009	0.009
0.25	1000	0.193	0.157	0.100	0.042	0.013
	2000	0.186	0.145	0.092	0.037	0.013
	5000	0.192	0.147	0.088	0.039	0.013
0.50	1000	0.216	0.208	0.178	0.110	0.047
	2000	0.215	0.205	0.168	0.103	0.045
	5000	0.214	0.198	0.178	0.102	0.043
1.00	1000	0.320	0.323	0.301	0.256	0.144
	2000	0.321	0.319	0.302	0.248	0.141
	5000	0.317	0.316	0.296	0.242	0.139
2.00	1000	0.560	0.560	0.547	0.510	0.383
	2000	0.556	0.556	0.541	0.504	0.379
	5000	0.543	0.544	0.537	0.504	0.374
5.00	1000	0.870	0.870	0.861	0.853	0.788
	2000	0.871	0.866	0.864	0.851	0.788
	5000	0.867	0.861	0.860	0.847	0.783

Table 2.7: MSE ratios of the estimate $\hat{W}_{XAB}$ of W given (X_1, A_X, B_X).

Table 2.7 gives the MSE ratio of $\hat{W}_{XAB}$, computed by simulation, for different parameter values. The continuous process X_t, $0 \le t \le 1$ was simulated by using 1000, 2000 and 5000 steps respectively for each parameter combination. We see that the MSE ratios obtained by using 2000 steps do not differ much from those obtained by using 5000 steps and therefore consider 2000 steps to be sufficient. The MSE ratios in Table 2.7 are consistently considerably smaller than those in Table 2.1, indicating the benefit of the extra information acquired through having available the extreme values of X_t in addition to X_1. The MSE ratios for $\iota = 2$ in Table 2.3 are larger, while those for $\iota = 10$ are slightly smaller, than their counterparts in Table 2.7. This indicates that the information gain through having available the extreme values of X_t corresponds to the information gain achieved in the intraday data situation of Section 2.5.3 with degrees of freedom ι somewhat less than 10, which we have verified to be around about 7. In Section 3.2 we shall model the return process on a stock by considering a duration of 1 day. In that framework, the accuracy gain achieved through having available the extreme values of X_t roughly corresponds to the accuracy gain achieved by having available 1-hourly intraday data.

For diagnostic purposes we are also interested in the distribution of $\hat{W}_{XAB}$. This distribution can be estimated empirically as accurately as required by generating enough realisations of $\hat{W}_{XAB}$ at given β and ψ, as follows:

- Generate a Brownian motion Z_t, $0 \le t \le 1$ and $W \sim UIG(\psi)$ (see Section 2.8).
- Set $X_t = \beta t W + \sqrt{W} Z_t$ and calculate $A_X = \max_{0 \le t \le 1}(X_t)$ and $B_X = -\min_{0 \le t \le 1}(X_t)$.
- Compute $\hat{W}_{XAB}$ from (2.7.40) with (x, a, b) replaced by (X_1, A_X, B_X).
- Repeat as many times as required.

2.8 Generating Independent Random Variables from the UIG Distribution

In Sections 2.3.5 and 2.5.3 we discussed methods for estimating W, Z and C from observations on X and V. Evaluating these methods require sampling from IG-related distributions, which we will discuss in this section. In (2.2.6) we have shown that, if W is $UIG(\psi)$-distributed, then for any $\delta, \gamma > 0$ subject to $\sqrt{\delta\gamma} = \psi$, $V = \delta W / \gamma$ is $IG(\delta, \gamma)$-distributed. Thus, to obtain an observation from an $IG(\delta, \gamma)$ distribution we need only multiply an observation from a $UIG\left(\sqrt{\delta\gamma}\right)$ distribution

by δ/γ. To obtain an algorithm for sampling from a $UIG(\psi)$ distribution we proceed as follows. Following Rydberg (1997) and Michael, Schucany and Haas (1976), from (2.2.2), we can write the UIG density function as

$$\begin{aligned} f_{UIG(\psi)}(w) &= \frac{\psi \exp(\psi^2)}{\sqrt{2\pi}} w^{-3/2} \exp\left(-\frac{1}{2}\psi^2\left(w^{-1}+w\right)\right) \\ &= \frac{\psi}{\sqrt{2\pi w^3}} \exp\left(-\frac{1}{2}\psi^2\left(w^{-1}+w-2\right)\right) \\ &= \frac{\psi}{\sqrt{2\pi w^3}} \exp\left(-\frac{1}{2}\left(\frac{\psi^2\left(w^2-2w+1\right)}{w}\right)\right) \\ &= \frac{\psi}{\sqrt{2\pi w^3}} \exp\left(-\frac{1}{2}\left(\frac{\psi^2(w-1)^2}{w}\right)\right). \end{aligned} \tag{2.8.1}$$

It then follows from Shuster (1968) that

$$B = \frac{\psi^2(W-1)^2}{W} \sim \chi_1^2 . \tag{2.8.2}$$

An observation of B, say b_0, can be easily generated as the square of a standard normal realisation. For each observation b_0 we must then solve (2.8.2) for w to obtain a corresponding observation from the UIG distribution. We rewrite (2.8.2) to get the quadratic equation

$$w^2 - \left(2+\frac{b_0}{\psi^2}\right)w + 1 = 0 . \tag{2.8.3}$$

This equation has roots (where w_1 is the smaller root) given by

$$\begin{aligned} w_1 &= \frac{1}{2}\left(2+b_0/\psi^2 - \sqrt{\left(2+b_0/\psi^2\right)^2-4}\right) \\ &= 1+\frac{b_0}{2\psi^2} - \frac{1}{2\psi^2}\sqrt{4b_0\psi^2+b_0^2} \\ w_2 &= \frac{1}{w_1} \end{aligned} \tag{2.8.4}$$

since the relationship which exists between the roots of any quadratic equation implies here that $w_1 w_2 = 1$. Michael, Schucany and Haas (1976) show that the smaller root w_1 should be chosen with probability $p_1(b_0)$ and the larger root should be chosen with probability $p_2(b_0)$, given by

$$p_1(b_0) = \frac{1}{1+w_1}$$
$$p_2(b_0) = 1 - p_1(b_0) \ . \tag{2.8.5}$$

Thus, the following algorithm can be used to sample an observation W from the $UIG(\psi)$ distribution:

- Generate $Z \sim N(0,1)$, $U \sim U(0,1)$ and $B \sim \chi_1^2$, all independently.
- Set $D = B/\psi^2$ and $w = \frac{1}{2}\left(2 + D - \sqrt{(2+D)^2 - 4}\right)$.
- With $I(A)$ the indicator of the event A, put $W = wI\left(U < 1/(1+w)\right) + (1/w)I\left(U \geq 1/(1+w)\right)$.

Atkinson (1982) presents an algorithm for generating $GIG(\lambda, \delta, \gamma)$ variates for any value of λ. This algorithm uses rejection with two-part envelopes. Once the parameters values for these envelopes have been determined, programming of the algorithm is simple and the efficiency (as measured by the proportion of prospective variates which are accepted) is reasonable for most values of λ, δ and γ.

Dagpunar (1989) argues that this algorithm is not easy to implement in totality because determination of the requisite envelope requires numerical optimisation of two parameters, and presents a short algorithm with efficiencies comparable to Atkinson (1982) but without the drawback of extensive numerical optimisation. This algorithm uses the ratio method (see Kinderman and Monahan (1977)) to generate prospective variates, which are then accepted or rejected according to a specific rule. We choose to use the algorithm we have presented above since it is easier to implement and does not contain acceptance/rejection rules.

The $UIG(\psi)$ distribution function is given in (2.2.3). Note that we could also generate an observation W from the $UIG(\psi)$ distribution by generating $U \sim U(0,1)$ and setting $W = F^{-1}_{UIG(\psi)}(U)$. To compute W we can solve the equation $F_{UIG(\psi)}(W) = U$ by using the Newton-Rhapson method. This algorithm has the advantage that it generates W's that are continuously dependent on the parameter ψ and the argument U. However, it does require about thirty times more computational effort than the algorithm described just below (2.8.5).

CHAPTER 3

GARCH MODELS AND BROWNIAN INVERSE GAUSSIAN PROCESSES

3.1 Introduction

In this Chapter we will use the theory of Chapter 2 to extend the normal GARCH model and contribute towards the literature as outlined in Section 1.3. These contributions will become clear as we discuss each of the various models we introduce. For ease of reference, Table 3.1 summarises all the models we will introduce and discuss in this chapter. The terms and symbols will become clear as we proceed to the sections which contains the equations indicated by the references column. Note the hierarchy in the models - every model can be seen as a special case of the ID-BIG model (the most complex model) which assumes an underlying BIG process for the innovations and caters for full intraday data. The HLC-BIG model follows from the ID-BIG model when we have only high, low and close returns available instead of full intraday data while the D-NIG model follows from the ID-BIG model when we have only daily returns available. In turn, the D-NOR, HLC-BM and ID-BM models follow from the D-NIG, HLC-BIG and ID-BIG models respectively when we assume normally distributed innovations instead of NIG-distributed innovations. Thus, although we talk about different models, we essentially have one model which we apply slightly differently according to the available data and the specification of the innovation distribution. It follows that these different models should all yield comparable parameter estimates, but we expect the models based on intraday data and assuming NIG-distributed innovations to lead to more accurate parameter estimates.

The remainder of this chapter is structured as follows. In Section 3.2.1 we state the daily normal-GARCH (D-NOR) model, fit it to generated and empirical data sets in Section 3.3 and, by means of residual analysis, find that an innovation distribution with heavier tails is required. The use of the so-called daily NIG-GARCH model (D-NIG) to allow for heavier tails is motivated in Section 3.2.2. We fit this model to the generated and empirical data sets and compare the results with those of the normal-GARCH model in Section 3.3. We find that the D-NIG model fits the data better than the D-NOR model. While this is not in itself a new contribution, we present it for the purpose of developing the ideas further in the subsequent sections.

As for new contributions, the daily NIG-GARCH model is extended to the ID-BIG model in Section 3.3.3, where we integrate intraday data and realised volatility into the NIG-GARCH context. We also discuss residual analysis to assess the quality of fit and consider some special cases of the

ID-BIG model, for example the ID-BM model where the underlying processes are Brownian motions. In addition, we note some volatility implications of the ID-BIG model and reflect on the relation between realised volatility, expected volatility and actual volatility. In Section 3.3.4 we extend the NIG-GARCH model to situations where high, low and close prices are available, instead of daily or full intraday data, leading to the HLC-BIG and HLC-BM models. Again, we discuss approaches to residual analysis, as well as technical difficulties hampering the implementation of some of these approaches.

In Section 3.3 we fit the models mentioned above to several data sets and discuss the results. The results are presented in terms of Quantile-Quantile (Q-Q) plots and Probability Integral Transform (PIT) plots as well as the Durbin-Watson (DW) test for first order correlation (introduced by Durbin and Watson (1951)) and the Q and Lagrange Multiplier (LM) tests for remaining ARCH effects (we review Q-Q and PIT plots, and motivate their use, in Section A.1 of the Appendix). We find that models assuming NIG innovations (or BIG processes) give more accurate parameter estimates than models assuming normal innovations (or underlying BM processes), and these estimates become more and more accurate as more and more data (higher intraday frequency) is used. We also illustrate and discuss the relation between realised volatility, the ordinary GARCH volatility estimator and the new adjusted volatility measure obtained from the model fit.

Finally, we find empirical evidence of time dependence in some of the model variables, which leads us to Chapter 4 and the further expansion of the ID-BIG model to allow for this dependence.

Data Case	Name	Model	Statistics	Special Cases	References
Daily Returns	D-NOR	$Y_n = \mu_n + \beta\sqrt{h_n} + \sqrt{h_n}Z_n$	Y_n	$\beta = 0$	(3.2.3)
	D-NIG	$Y_n = \mu_n + \beta\sqrt{h_n}W_n + \sqrt{h_n W_n}Z_n$			(3.2.7)
Intraday Returns	ID-BM	$Y_{t,n} = \left(\mu_n + \beta\sqrt{h_n}\right)t + \sqrt{h_n}Z_{t,n}$	$Y_n \equiv Y_{1,n}$, S_n	$\beta = 0$ 5/10/30 minutes	(3.2.10)
	ID-BIG	$Y_{t,n} = \left(\mu_n + \beta\sqrt{h_n}W_n\right)t + \sqrt{h_n W_n}Z_{t,n}$			(3.2.12)
High-Low-Close Returns	HLC-BM	$Y_{t,n} = \left(\mu_n + \beta\sqrt{h_n}\right)t + \sqrt{h_n}Z_{t,n}$	$Y_n \equiv Y_{1,n}$, A_n, B_n	$\beta = 0$	(3.2.22)
	HLC-BIG	$Y_{t,n} = \left(\mu_n + \beta\sqrt{h_n}W_n\right)t + \sqrt{h_n W_n}Z_{t,n}$			(3.2.23)

Table 3.1: Overview of models introduced in Chapter 3.

3.2 Models

3.2.1 The Normal-GARCH Model for Daily Returns

Let $p_1, p_2, \ldots$ denote the daily closing prices of a financial instrument and write $Y_n = \log(p_n/p_{n-1})$ for the return on day n. A GARCH model to describe the return series takes the form

$$Y_n = \mu_n + \sqrt{h_n} X_n \text{ for } n = 1, 2, \ldots \tag{3.2.1}$$

where μ_n represents an expected (or structural) component, h_n is the volatility and X_n the innovation at time n. It is assumed that μ_n and h_n are at most dependent on the past information set F_{n-1} (i.e. the σ-field generated by $Y_1, Y_2, \ldots, Y_{n-1}$ or a larger dataset available at time $n-1$). Typical choices are $\mu_n = \nu + \phi Y_{n-1}$ and $h_n = \alpha_0 + \alpha_1 (Y_{n-1} - \mu_{n-1})^2 + \beta_1 h_{n-1}$ which yield the popular AR(1)-GARCH(1,1) model. The innovations X_n are supposed to be independent of F_{n-1} and also independently and identically distributed (iid). We denote their common distribution function by $G(x)$ and their density by $g(x)$. The log-likelihood function of the data $Y_1, \ldots, Y_N$ (assuming N days) may then be written as

$$\sum_{n=1}^{N} \left(\log g\left(\frac{Y_n - \mu_n}{\sqrt{h_n}} \right) - \frac{1}{2} \log h_n \right). \tag{3.2.2}$$

Often, the common distribution of the innovations is assumed to be the $N(\beta, 1)$ distribution, i.e. $X_n = \beta + Z_n$, in which case (3.2.1) may be written as

$$Y_n = \mu_n + \beta\sqrt{h_n} + \sqrt{h_n} Z_n \text{ for } n = 1, 2, \ldots \tag{3.2.3}$$

where $Z_1, Z_2, \ldots$ are iid $N(0,1)$-distributed. This is commonly referred to as a GARCH-in-mean model. We will refer to model (3.2.3) as the D-NOR model, where D refers to the case of daily returns data and NOR refers to the normally distributed innovations.

The log-likelihood function (3.2.2) for model (3.2.3) then becomes

$$\sum_{n=1}^{N} \left(\log \varphi\left(\frac{Y_n - \mu_n}{\sqrt{h_n}} - \beta \right) - \frac{1}{2} \log h_n \right)$$

$$= \sum_{n=1}^{N} \left(\log \left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{Y_n - \mu_n}{\sqrt{h_n}} - \beta \right)^2 \right) \right) - \frac{1}{2} \log h_n \right)$$

$$= -\frac{N}{2}\log 2\pi - \frac{1}{2}\sum_{n=1}^{N}\left(\left(\frac{Y_n-\mu_n}{\sqrt{h_n}}-\beta\right)^2 - \log h_n\right). \qquad (3.2.4)$$

(3.2.4) depends upon the parameters occurring in the specification of μ_n and h_n. Therefore, in order to proceed with maximum likelihood inference, we must now specify μ_n and h_n. Since μ_n is generally quite small, a popular choice is $\mu_n = 0$, for example when analysing currency exchange return series (see Forsberg and Bollerslev (2002)). Other popular choices include $\mu_n = \nu + \phi Y_{n-1}$ (the AR(1) model) and more general autoregressive moving average (ARMA) models. A huge number of specifications for h_n have been suggested in the GARCH literature. Hansen and Lunde (2001) find that a GARCH(1,1) model is adequate in most cases. Combined with the AR(1) specification on μ_n, this means that

$$\begin{aligned} \mu_n &= \nu + \phi Y_{n-1} \\ h_n &= \alpha_0 + \alpha_1 (Y_{n-1}-\mu_{n-1})^2 + \beta_1 h_{n-1}\ , \end{aligned} \qquad (3.2.5)$$

where $|\phi|<1$, $\alpha_0>0$, $0<\alpha_1<1$, $0<\beta_1<1$ and $0<\alpha_1+\beta_1<1$.

Model (3.2.3) with the GARCH(1,1) specification for h_n corresponds to the non-linear asymmetric GARCH (NAGARCH(1,1)) specification of Engle and Ng (1993). From (3.2.3) it follows that $Y_{n-1}-\mu_{n-1} = \beta\sqrt{h_{n-1}} + \sqrt{h_{n-1}}Z_{n-1}$. Setting $\varepsilon_{n-1} = \sqrt{h_{n-1}}Z_{n-1}$ we see that $(Y_{n-1}-\mu_{n-1})^2 = \left(\varepsilon_{n-1}+\beta\sqrt{h_{n-1}}\right)^2$, so that the "news impact curve" is symmetric around $\varepsilon_{n-1} = -\beta\sqrt{h_{n-1}}$ (see Engle and Ng (1993)).

Even if the underlying innovation distribution is not conditionally normal, the ML estimates based on the normality assumption (i.e. Pseudo Maximum Likelihood (PML) or Quasi Maximum Likelihood (QML) estimates) are asymptotically consistent and have an asymptotically normal distribution (see Bollerslev and Wooldridge (1992) and Gouriéroux (1997)).

Residual analysis

It follows from (3.2.3) that $Z_n = (Y_n-\mu_n)/\sqrt{h_n} - \beta$. Thus, the corresponding residuals $\hat{Z}_n = (Y_n-\hat{\mu}_n)/\sqrt{\hat{h}_n} - \hat{\beta}$ (calculated using the estimators $\hat{\mu}_n$ and $\hat{h}_n$ obtained from the maximum likelihood fit of (3.2.4)) should be approximately $N(0,1)$-distributed, while $\hat{X}_n = \hat{\beta} + \hat{Z}_n$ should be

approximately $N(\hat{\beta},1)$-distributed. We can test the normality assumptions by plotting the order statistics of the residuals against those of the normal distribution by means of a Q-Q plot and a PIT plot. Of course, the tests on $\hat{Z}_n$ and $\hat{X}_n = \hat{\beta} + \hat{Z}_n$ will yield identical results since $\hat{X}_n$ just shifts $\hat{Z}_n$ by the value $\hat{\beta}$. We indicate these results by using the caption $X = Z$ in the tables and figures in Section 3.3. Many authors (see e.g. Forsberg (2002)) have found that an innovation distribution with heavier tails is required. We illustrate the fit and residual analysis of model (3.2.3) and confirm the need for innovation distributions with heavier tails in Section 3.3.

3.2.2 The NIG-GARCH Model for Daily Returns

Recently a number of authors have argued in favour of using the NIG distribution as an innovation distribution with heavier tails in model (3.2.1) (see e.g. Andersson (2001), Jensen and Lunde (2001), Venter and de Jongh (2002, 2004), Forsberg (2002) and especially the important paper of Forsberg and Bollerslev (2002)). This has lead to what may be termed the NIG-GARCH model, which may be formulated as follows using the notions introduced in Chapter 2. Let $W_1, W_2, \ldots$ be iid $UIG(\psi)$-distributed independently of $Z_1, Z_2, \ldots$ and take $X_n = \sqrt{W_n} Z_n$ so that (3.2.1) may be written as

$$Y_n = \mu_n + \sqrt{h_n}\left(\sqrt{W_n} Z_n\right) = \mu_n + \sqrt{h_n W_n} Z_n \ . \tag{3.2.6}$$

From Section 2.3 we know that $\sqrt{W_n} Z_n$ has a $SNIG(0,\psi)$-distribution, so that the first expression in (3.2.6) differs from the model (3.2.3) in the replacement of the $N(0,1)$-distributed innovation by a $SNIG(0,\psi)$-distributed innovation. The second expression in (3.2.6) indicates that we may interpret this model alternatively as keeping the $N(0,1)$-distributed innovation Z_n but adjusting the GARCH volatility h_n by the random factor W_n to a new volatility measure $h_n W_n$, which we will now interpret and argue is a better indication of the actual volatility that will be experienced on day n. The GARCH volatility measure h_n depends upon information up to close of market on day $n-1$ and therefore cannot incorporate overnight effects on prices due to news arriving after close of market or the effect of other markets around the world. Mostly, no tangible observations of these effects are available so we model them by random impact factors W_n which adjust the GARCH volatility h_n to a new volatility measure $h_n W_n$. The assumption $W_n \sim UIG(\psi)$ entails that

$E(W_n)=1$; thus there is no adjustment on average but on day n the adjustment may be up or down as reflected by the value of W_n. Since $E(h_nW_n|\mathsf{F}_{n-1})=h_n$, the GARCH volatility h_n represents the **expected volatility** on day n (given past data) while h_nW_n represents the **actual volatility**. This relation will become more apparent in Section 3.3 where we will consider the empirical results after fitting the model to generated and empirical data.

More generally, we could take $X_n=\beta W_n+\sqrt{W_n}Z_n$ so that (3.2.6) is generalised to

$$Y_n=\mu_n+\sqrt{h_n}\left(\beta W_n+\sqrt{W_n}Z_n\right)=\mu_n+\beta\sqrt{h_n}W_n+\sqrt{h_nW_n}Z_n\ , \qquad (3.2.7)$$

thus allowing for a possible shift adjustment of the impact factor on the structural part of the model also. We will indicate this model by D-NIG, where D indicates that we are modelling daily returns and NIG refers to the normal inverse Gaussian distributed innovations, since we know from (2.3.2) that $X_n=\beta W_n+\sqrt{W_n}Z_n$ is $SNIG(\beta,\psi)$-distributed. This model is comparable to a GARCH-in-mean model which would have the term $\beta\sqrt{h_n}W_n$ replaced by $\beta\sqrt{h_n}$. Again, since $E(W_n)=1$, on average $\sqrt{h_n}$ and $\sqrt{h_n}W_n$ are equal but one can argue that if volatility also affects the structural part of the model then allowance should be made to incorporate also the random impact W_n due to effects of factors operating after close of market on day $n-1$. Then (3.2.7) should yield a more realistic GARCH-in-mean formulation.

Model (3.2.7) with the GARCH(1,1) specification for h_n amounts to a refinement of the NAGARCH(1,1) specification of Engle and Ng (1993). The NAGARCH(1,1) specification and news impact curve is discussed in Section 3.1.1 above, just below (3.2.3).

If, in the present situation, we write $\varepsilon_{n-1}=\sqrt{h_{n-1}W_{n-1}}Z_{n-1}$ for the error term in Y_{n-1}, (3.2.7) becomes

$$Y_{n-1}-\mu_{n-1}=\beta\sqrt{h_{n-1}}W_{n-1}+\varepsilon_{n-1}\ . \qquad (3.2.8)$$

Then the news impact curve follows from $\left(Y_{n-1}-\mu_{n-1}\right)^2=\left(\varepsilon_{n-1}+\beta\sqrt{h_{n-1}}W_{n-1}\right)^2$ and has the NAGARCH(1,1) form except that its minimum is at the random location $\varepsilon_{n-1}=-\beta\sqrt{h_{n-1}}W_{n-1}$ rather than the fixed (given the past information F_{n-1}) point $\varepsilon_{n-1}=-\beta\sqrt{h_{n-1}}$. Again this refinement appears appropriate to allow for the impact of news occurring after the close of market on day $n-1$.

Since $X_n = \beta W_n + \sqrt{W_n} Z_n$ is $SNIG(\beta, \psi)$-distributed, it follows from (3.2.2) that the log-likelihood function for model (3.2.7) is

$$\sum_{n=1}^{N} \left(\log f_{SNIG(\beta,\psi)} \left(\frac{Y_n - \mu_n}{\sqrt{h_n}} \right) - \frac{1}{2} \log h_n \right) . \tag{3.2.9}$$

Maximum likelihood estimation (MLE) may then be used to obtain estimates for the parameters of model (3.2.7). We use model (3.2.7) with the AR(1)-GARCH(1,1) specification by default unless otherwise indicated.

Residual analysis

Once we have estimates $\hat{\beta}$, $\hat{\psi}$, $\hat{\mu}_n$ and $\hat{h}_n$ we can calculate the residuals $\hat{X}_n = (Y_n - \hat{\mu}_n)/\sqrt{\hat{h}_n}$ and test them against their implied $SNIG(\hat{\beta}, \hat{\psi})$ distribution by means of Q-Q plots and PIT plots. Further checks can be based on estimates of the latent variables in the model. We can estimate W_n and Z_n by $\hat{W}_{X,n}$ and $\hat{Z}_{X,n}$, which is computed by (2.3.18) and (2.3.21) with X replaced by $\hat{X}_n$, and then compare the $\hat{W}_{X,n}$'s and $\hat{Z}_{X,n}$'s against their model implied distributions computed as indicated at the end of Section 2.3.5. We illustrate this in Section 3.3.

Finally, note that the results obtained for the D-NOR model (3.2.3) of Section 3.2.1 follows as a special case of the results for the D-NIG model (3.2.7), if all the W_n's are taken to be 1, which is equivalent to setting $\psi = \infty$.

3.2.3 GARCH Models Based on Intraday BIG Processes

Traditionally, GARCH models were fitted to daily data and gave daily volatility estimates. In recent times intraday data (e.g. 5 minute and tick-by-tick data) have become more and more obtainable and reliable. Andersen et al. (2003) states that "forecasting performance has improved with the incorporation of more data, not only since high-frequency volatilities turn out to be highly predictable, but also because the information in high-frequency data proves useful for forecasting at longer horizons." An entire field of study has developed in which the microstructure of financial time series is analysed. However, progress in volatility modelling has slowed in some respects in the last few years due to a number of difficulties. According to Andersen et al. (2003), "standard volatility models used for forecasting at the daily level cannot readily accommodate the information in intraday data, and models specified directly for the intraday data generally fail to capture the

longer interdaily volatility movements sufficiently well." As a result, standard practice is still to produce forecasts of daily volatility from daily return observations, even though intraday data may be available.

Set against this background, Andersen et al. (2003) propose a new framework for volatility forecasting by focusing on an empirical measure of daily return variability called realised volatility (RV), which is easily computed from intraday returns. This framework produces significant improvements in predictive performance compared to standard procedures that rely on daily data alone by efficiently exploiting the information in intraday return data without having to explicitly model the intraday data. At this point a question arises, namely what the relation is between the volatility measures produced by GARCH models and those produced by realised volatility. We will clarify this relation by extending the NIG-GARCH model of Section 3.1.2 to cater for intraday data and in the process showing that the actual daily volatility $h_n W_n$ as introduced in Section 3.1.2 closely approximates the realised volatility. This requires some additional notation.

Extending the notation of Sections 1.2 and 2.7.3, assume intraday prices are available and let $p_{t,n}$ denote the price at time t on day n where we measure time in terms of a fraction of the trading day (so that $0 \le t \le 1$, $t=0$ corresponds to opening and $t=1$ to closing time). Also let $Y_{t,n} = \log\left(p_{t,n}/p_{0,n}\right)$ denote the return up to time t on day n. Further, break up the trading day into I periods of equal length $\Delta = 1/I$ each and assume that we have the prices at least at the times Δi, $i=1,2,\ldots,I$. Let $D_{i,n} = \log\left(p_{i\Delta,n}/p_{(i-1)\Delta,n}\right) = Y_{i\Delta,n} - Y_{(i-1)\Delta,n}$ denote the return over the i-th period, $i=1,2,\ldots,I$. Then on the n-th day the total return is $Y_{1,n} \equiv Y_n = \sum_{i=1}^{I} D_{i,n}$ and the corresponding realised volatility is defined by $R_n = \sum_{i=1}^{I} D_{i,n}^2$. Here we introduce two models for the intraday return process $Y_{t,n}$, $0 \le t \le 1$, which are extensions of the models (3.2.3) and (3.2.7) and will enable us to study the important advantages brought about by having available intraday prices in addition to closing prices.

Let $Z_{t,n}$ be standard Brownian motions in t, independent over n. For a first model, we follow and extend the work of Lildholdt (2002), modelling the intraday returns by the BM process

$$Y_{t,n} = \left(\mu_n + \beta\sqrt{h_n}\right)t + \sqrt{h_n}Z_{t,n} \ . \qquad (3.2.10)$$

We will denote this model by ID-BM, where ID indicates that the model caters for intraday data and BM indicates the use of Brownian motion. Here μ_n and h_n are again assumed at most dependent

on the past information set F_{n-1} (now the σ-field generated by $Y_{t,m}$, $0 \le t \le 1$, $m < n$). Of course the choices of μ_n and h_n must still be stated to make (3.2.10) a fully specified intraday returns process model. We now have to derive an expression for the log-likelihood function. Since model (3.2.10) is a special case of the next model we will introduce, we postpone the derivation of the log-likelihood function until later in this section.

As a second and more complicated model, we generalise the ID-BM model by replacing the BM process with a BIG process, as follows. Let W_n be iid $UIG(\psi)$-distributed and $Z_{t,n}$ be standard Brownian motions in t, independent over n, and also independent of the W_n's. From Section 2.7.2 it follows that each

$$X_{t,n} = \beta t W_n + \sqrt{W_n} Z_{t,n}, \quad 0 \le t \le 1 , \tag{3.2.11}$$

follows a SBIG process in t, independently over n. Further extending the work of Lildholdt (2002), we model the intraday returns by the BIG process

$$Y_{t,n} = t\mu_n + \sqrt{h_n} X_{t,n} = \left(\mu_n + \beta\sqrt{h_n} W_n\right) t + \sqrt{h_n W_n} Z_{t,n} . \tag{3.2.12}$$

Whichever GARCH-type specification of μ_n and h_n is used, we will refer to the model resulting from (3.2.12) as the ID-BIG model, where ID indicates intraday data and BIG indicates the Brownian Inverse Gaussian processes. We now derive an expression for the log-likelihood function which is valid for any specification of μ_n and h_n. We take the periodic intraday returns $D_{i,n}$, $i = 1, 2, \ldots, I$ as our observed data. From Section 2.7.3 we see that Y_n and $R_n = \sum_{i=1}^{I} D_{i,n}^2$ constitute sufficient statistics for $D_{i,n}$, $i = 1, 2, \ldots, I$, and from (2.7.21) the joint conditional density of Y_n and $\tilde{V}_n = IS_n$, with $S_n = R_n - Y_n^2 / I$, given F_{n-1} is given by

$$f_{Y_n, \tilde{V}_n}(y_n, \tilde{v}_n) = h_n^{-3/2} f_{SNCIG(\beta, \psi, I-1)}\left(\frac{y_n - \mu_n}{h_n^{1/2}}, \frac{\tilde{v}_n}{h_n}\right) . \tag{3.2.13}$$

Hence, the log-likelihood function of the $\left(Y_n, \tilde{V}_n\right)$, $n = 1, \ldots, N$ data may be written as

$$\sum_{n=1}^{N}\left(\log f_{SNCIG(\beta, \psi, I-1)}\left(\frac{Y_n - \mu_n}{h_n^{1/2}}, \frac{\tilde{V}_n}{h_n}\right) - \frac{3}{2}\log h_n\right) \tag{3.2.14}$$

and standard maximum likelihood inference may be carried out. We developed SAS® (based on Proc NLP) and FORTRAN® programs to compute ML estimates and corresponding asymptotic

standard errors. This was a rather delicate task and we give more details in Section A.3 of the Appendix.

Residual analysis

Once we have estimates $\hat{\beta}, \hat{\psi}, \hat{\mu}_n$ and $\hat{h}_n$ we can do a number of diagnostic checks:

- Firstly, as in Section 3.1.2, $\hat{X}_n \equiv (Y_n - \hat{\mu}_n)/\sqrt{\hat{h}_n}$ should be approximately $SNIG(\hat{\beta}, \hat{\psi})$-distributed.
- Secondly, from (2.7.20) of Section 2.7.3, identifying δ^2 with h_n, $V_n \equiv IS_n/h_n = W_n C_n$ (with C_n as in (2.7.16)) is $SCIG(\psi, I-1)$-distributed and therefore $\hat{V}_n \equiv IS_n/\hat{h}_n$ should be approximately $SCIG(\hat{\psi}, I-1)$-distributed.
- Thirdly, it follows from (2.6.1) that

$$T_n = \frac{X_{1,n}}{\sqrt{V_n/(I-1)}} = \left(\frac{Y_n - \mu_n}{\sqrt{h_n}}\right) \Big/ \sqrt{\frac{IS_n/h_n}{I-1}} = \sqrt{\frac{I-1}{I}}\,\frac{Y_n - \mu_n}{\sqrt{S_n}} \qquad (3.1.15)$$

is $TIG(\beta, \psi, I-1)$-distributed. Hence $\hat{T}_n = \sqrt{I-1}(Y_n - \hat{\mu}_n)/\sqrt{IS_n}$ should be approximately $TIG(\hat{\beta}, \hat{\psi}, I-1)$-distributed.

Further checks can be based on estimates of the latent variables in the model. We can estimate W_n by $\hat{W}_{XV,n}$ which is computed by (2.5.8) with X and V replaced by $\hat{X}_n$ and $\hat{V}_n$, and then compare the $\hat{W}_{XV,n}$'s against their model-implied distribution computed as indicated at the end of Section 2.5.3. An alternative here would be to use the simpler estimator $\hat{V}_n/(I-1)$ instead of $\hat{W}_{XV,n}$. Analogous estimates for $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ can be calculated and compared against their respective model distributions.

Special cases

Note that the choice $\psi = \infty$ forces all the W_n's to take the value 1 and reduces the ID-BIG model (3.2.12) to the ID-BM model (3.2.10), for which we now derive the log-likelihood function. In this case (2.7.18) and (2.7.20) show that $(Y_n - \mu_n)/\sqrt{h_n}$ is $N(\beta, 1)$-distributed independent of $V_n = IS_n/h_n$ which is χ^2_{I-1}-distributed (we may say that the pair is "normal chi-square" distributed). Hence, for $\psi = \infty$ the log-likelihood function of the (Y_n, S_n) data may be written as

$$\sum_{n=1}^{N}\left(\log\varphi\left(\frac{Y_n-\mu_n}{\sqrt{h_n}}-\beta\right)+\log f_{\chi^2_{I-1}}\left(\frac{IS_n}{h_n}\right)-\frac{3}{2}\log h_n+\log(I)\right). \tag{3.2.16}$$

If we take $\beta=0$ in addition to $\psi=\infty$ then (3.2.12) simply becomes $Y_{t,n}=t\mu_n+\sqrt{h_n}Z_{t,n}$, which is essentially the model treated by Lildholdt (2002). Thus (3.2.12) generalises the model of Lildholdt (2002) and the ID-BM model to the ID-BIG model.

There are some more special cases we could examine. For instance, it follows from (3.2.2) that the log-likelihood function using only the realised volatility data, or equivalently only the SSD data S_n, is

$$\sum_{n=1}^{N}\left(\log f_{SCIG(\psi,I-1)}\left(\frac{IS_n}{h_n}\right)-\log h_n+\log(I)\right). \tag{3.2.17}$$

We could maximise this function separately to get further estimates of the parameters. Notice that (3.2.17) does not involve the parameter β so that no estimate of it will be forthcoming from maximising (3.2.17). Also, μ_n appears only indirectly in (3.2.17) through the specification of h_n so that parameters involved in μ_n may not be estimated well from (3.2.17). Of course, estimates obtained in this way should agree with the full ML estimates obtained from (3.2.14). We applied this model to our generated data sets and found that this is indeed the case. The results do not produce any significant additional insights. Furthermore, it is highly unlikely to have only the daily realised volatilities available, and not the daily returns also. Thus, in practice, we would rather fit the complete model that utilises the return and realised volatility data. Therefore we do not report on this possibility further.

Returning to the general case, we note some volatility implications of the ID-BIG model. As noted in Section 3.2.2 above the actual volatility to be experienced on day n is represented by h_nW_n. It follows from (2.7.19) that

$$S_n=\tilde{V}_n/I=h_nW_nC_n/I\ . \tag{3.2.18}$$

When I is large, C_n/I should be about 1, by the law of large numbers, so that $S_n\approx h_nW_n$ and S_n may be thought of as estimate of the actual volatility on day n. In as much as R_n and S_n are typically close to each other, the same holds for the realised volatility R_n, i.e. we should have $R_n\approx h_nW_n$. This is in line with results in the realised volatility literature (see e.g. Andersen et al. (2000 and 2001)). We can check this approximation in practice, since fitting the ID-BIG model

leads to the estimates $\hat{h}_n$ and $\hat{W}_n$ and hence also to the alternative estimator $\hat{h}_n\hat{W}_n$ of the actual volatility h_nW_n. We should then have $R_n \approx \hat{h}_n\hat{W}_n$ if I is large. We shall verify that this is the case in an example in Section 3.3.

We can also look at (3.2.18) somewhat differently. Writing it in the form $S_n/h_n = W_nC_n/I \approx W_n$, it follows that $S_n/\hat{h}_n$ should be approximately $UIG(\hat{\psi})$-distributed if I is large. Forsberg and Bollerslev (2002) present empirical evidence to this effect. We also present such empirical evidence in Section 3.3 where we analyse empirical data. This is closely related to the diagnostic check based on $\hat{V}_n$ which has the added advantage that it does not require I to be large.

3.2.4 GARCH Models Based on High, Low and Close Prices

We discussed models based on close prices alone in Sections 3.2.1 and 3.2.2 and models based on intraday data in Section 3.2.3, and we expect the latter to provide more accurate estimates of the relevant parameters since they use more data. However, the availability of intraday data is a relatively recent phenomenon. Because of the extensive effort required to gather tick-by-tick data, intraday prices at regular intervals may not be readily available. A situation that arises more frequently is that returns of financial instruments are summarised in terms of high, low and close prices. In such cases, where intraday data is unavailable and we therefore cannot fit models based on intraday data, we expect to obtain better models when we use high and low prices in addition to close prices alone. Thus, we have an intermediate option between using only daily close prices and using full intraday data. We will now extend the models of Sections 3.2.1 and 3.2.2 to cater for high and low prices also. This requires some notation in addition to that of Sections 3.2.1 to 3.2.3.

Let p_n^o, p_n^h, p_n^l and p_n^c denote the open, high, low and close prices of a financial instrument on day n. Note that $p_n^o = p_{n-1}^c$, so that we actually need only the high, low and close prices. Again, following and extending the work of Lildholdt (2002), and assuming the intraday returns follow the BM process (3.2.10), the total return on day n is then

$$Y_n = \log\left(p_n^c\right) - \log\left(p_n^o\right) = Y_{1,n} = \mu_n + \beta\sqrt{h_n} + \sqrt{h_n}Z_{1,n} \ . \tag{3.2.19}$$

The high price on day n may be transformed to the high return

$$
\begin{aligned}
A_n &= \log\left(p_n^h\right) - \log\left(p_n^o\right) \\
&= \max_{0\le t\le 1}\left(\log\left(p_{t,n}\right) - \log\left(p_{0,n}\right)\right) \\
&= \max_{0\le t\le 1}\left(Y_{t,n}\right)
\end{aligned} \tag{3.2.20}
$$

and the low price on day n may be transformed to the low return

$$
\begin{aligned}
B_n &= \log\left(p_n^o\right) - \log\left(p_n^l\right) \\
&= -\min_{0\le t\le 1}\left(\log\left(p_{t,n}\right) - \log\left(p_{0,n}\right)\right) \\
&= -\min_{0\le t\le 1}\left(Y_{t,n}\right) .
\end{aligned} \tag{3.2.21}
$$

Given the past information set F_{n-1}, μ_n and h_n can be calculated. The joint density function of the $\left(Y_n, A_n, B_n\right)$ data is given by (2.7.32), so that the log-likelihood function of the $\left(Y_n, A_n, B_n\right)$ data may now be written in terms of (2.7.32) as

$$
\sum_{n=1}^{N} \log f_{U_1, A_U, B_U}\left(Y_n, A_n, B_n \middle| \mu_n, h_n\right) . \tag{3.2.22}
$$

We shall indicate this model by HLC-BM, where HLC indicates high, low and close data and BM indicates Brownian motion.

As in Section 3.2.3 we can generalise the HLC-BM model by replacing the BM process (3.2.10) with a BIG process (3.2.12). In this case the joint density function of the $\left(Y_n, A_n, B_n\right)$ data is given by (2.7.34), and we may write the log-likelihood function of the $\left(Y_n, A_n, B_n\right)$ data in terms of (2.7.34) as

$$
\sum_{n=1}^{N} \log f_{Y,A,B}\left(Y_n, A_n, B_n \middle| \mu_n, h_n\right) . \tag{3.2.23}
$$

We shall indicate this model by HLC-BIG, where HLC again indicates high, low and close prices and BIG indicates Brownian inverse Gaussian processes. Again, standard maximum likelihood inference may now be used to fit this model. In Section A.3 of the Appendix we give more details on the SAS® (based on Proc NLP) and FORTRAN® programs we developed to compute ML estimates and corresponding asymptotic standard errors. Note that, analogous to Sections 3.2.1 to 3.2.3, the HLC-BM model is a special case of the HLC-BIG process, obtained from letting all the W_n's take on the value 1 or, equivalently, setting $\psi = \infty$.

Residual analysis

As to model checking, we can again compute total return residuals $\hat{X}_n = (Y_n - \hat{\mu}_n)/\sqrt{\hat{h}_n}$ and compare them against a $SNIG(\hat{\beta}, \hat{\psi})$-distribution as in Section 3.2.2. Also, we can compute the estimates $\hat{W}_{X,n}$ and $\hat{Z}_{X,n}$ as in Section 3.2.2 and use them as before. However, these residual estimates do not use the high and low returns and in principle we expect to do better if we use that data as well. We now show how residual analysis may be approached using high and low returns in addition to daily returns and we indicate technical difficulties encountered along the way. Supposing we have the full intraday returns process data $Y_{t,n}$, $0 \le t \le 1$, we can introduce the residual process $\hat{X}_{t,n} = (Y_{t,n} - t\hat{\mu}_n)/\sqrt{\hat{h}_n}$, $0 \le t \le 1$, which approximates the process $X_{t,n} = \beta t W_n + \sqrt{W_n} Z_{t,n}$, $0 \le t \le 1$. Further define $\hat{A}_{X,n} = \max_{0 \le t \le 1}(\hat{X}_{t,n})$ which approximates $A_{X,n} = \max_{0 \le t \le 1}(X_{t,n})$ and let $\hat{B}_{X,n} = -\min_{0 \le t \le 1}(\hat{X}_{t,n})$ which approximates $B_{X,n} = -\min_{0 \le t \le 1}(X_{t,n})$. Then the joint distribution of $(\hat{X}_{1,n}, \hat{A}_{X,n}, \hat{B}_{X,n})$ would be approximately that of $(X_{1,n}, A_{X,n}, B_{X,n})$. This suggests that we can estimate W_n by $\hat{W}_{XAB,n}$ which is computed by (2.7.40) with (X_1, A_X, B_X) replaced by $(\hat{X}_{1,n}, \hat{A}_{X,n}, \hat{B}_{X,n})$, and then compare the $\hat{W}_{XAB,n}$'s against their model-implied distribution computed as indicated at the end of Section 2.7.5.

Unfortunately we need the full intraday returns process data $Y_{t,n}$, $0 \le t \le 1$ to be able to calculate $\hat{A}_{X,n}$ and $\hat{B}_{X,n}$, and this is typically not available. In general there is no way to calculate $\hat{A}_{X,n}$ and $\hat{B}_{X,n}$ directly from A_n and B_n (or only the available high, low and close prices). Thus, in the absence of full intraday data this residual analysis approach is generally not feasible. If full intraday data are indeed available, it will be better to proceed according to Section 3.2.3. One special case where this residual analysis approach is indeed feasible is when we may assume that $\mu_n = 0$. Then $\hat{\mu}_n = 0$, $\hat{X}_{t,n} = Y_{t,n}/\sqrt{\hat{h}_n}$, $\hat{A}_{X,n} = A_n/\sqrt{\hat{h}_n}$, $\hat{B}_{X,n} = B_n/\sqrt{\hat{h}_n}$ and we do not need any further data to be able to proceed as above. We have used this approach to estimate $\hat{W}_{XAB,n}$ and the results are given in Section 3.3. However, even in this case a remaining concern is that the implied computational burden is quite heavy.

3.3 Empirical Illustrations

In this section we illustrate the results of the models presented above in terms of the following three data sets:

- GenBM: We generate 1000 daily returns from an AR(1)-GARCH(1,1) model with parameters $\nu = 0$, $\phi = 0.15$, $\alpha_0 = 10^{-5}$, $\alpha_1 = 0.1$, $\beta_1 = 0.8$ and $\beta = 0.2$ where $Y_{t,n} = \left(\mu_n + \beta\sqrt{h_n}\right)t + \sqrt{h_n}Z_{t,n}$ is assumed to follow a continuous BM process (model (3.2.10) of Section 3.2.3). We have seen at the end of Section 2.7.5 that the continuous BM process leading to the daily high, low and close prices can be approximated well by using 2000 intraday points. In this case we use 2028 intraday points, since 2028 is divisible by 78, 39 and 13, which corresponds to 5, 10 and 30 minute intervals. We record the daily high, low and total returns, as well as the realised volatilities corresponding to the 5, 10 and 30 minute intervals. We also allow an extra 50 starting days to allow μ_n and h_n to stabilise.
- GenBIG: We generate 1000 daily returns by following the same procedure (with the same parameter values) as with the GenBM data, except that $Y_{t,n} = \left(\mu_n + \beta\sqrt{h_n W_n}\right)t + \sqrt{h_n W_n}Z_{t,n}$ is assumed to follow a BIG process (model (3.2.12) of Section 3.2.3), and where W_n is iid $UIG(\psi)$-distributed with $\psi = 2$. What is more, we use the same underlying $Z_{t,n}$'s, so that the results of fitting the various models (see below) to the GenBM and GenBIG data sets can be compared.
- IBM: We use intraday data of the IBM share return series obtained from Tick Data, Inc. and Price Data, Inc. for 1000 (plus 50 days to allow μ_n and h_n to stabilise) consecutive days over the period 1999/02/09 to 2003/01/31. We use daily return data as well as high and low return data and also compute realised volatilities corresponding to 5 minute, 10 minute and 30 minute intraday periods.

The parameter choices for the generated data sets are suggested by typical values obtained for these parameters when the corresponding models are fitted to price series of shares traded on the Johannesburg Stock Exchange (JSE), for example Anglo American and Sasol. We use the generated data sets to illustrate the theory under known parameter values. This also serves as a test for our programs and derivations.

We fit AR(1)-GARCH(1,1) models, i.e. we use the specifications $\mu_n = \nu + \phi Y_{n-1}$ and $h_n = \alpha_0 + \alpha_1\left(Y_{n-1} - \mu_{n-1}\right)^2 + \beta_1 h_{n-1}$ explained above. The models, summarised earlier in Table 3.1, are abbreviated as follows:

- D-NOR: using the total daily return data Y_n only, assuming normally distributed innovations and maximising (3.2.4).
- D-NIG: using the total daily return data Y_n only, assuming $NIG(\beta,\psi)$-distributed innovations and maximising (3.2.9).
- HLC-BM: using the total, high and low return data (Y_n, A_n, B_n), assuming an underlying BM process and maximising (3.2.22).
- HLC-BIG: using the total, high and low return data (Y_n, A_n, B_n), assuming an underlying BIG process and maximising (3.2.23).
- ID-BM5/10/30: using daily return data as well as 5/10/30 minute $(I = 78/39/13)$ realised volatility data, assuming an underlying BM process and maximising (3.2.16).
- ID-BIG5/10/30: using daily return data as well as 5/10/30 minute $(I = 78/39/13)$ realised volatility data, assuming an underlying BIG process and maximising (3.2.14).

The quality of the fits are presented in terms of Q-Q plots and PIT plots (see Section A.1 of the Appendix), as well as the Durbin-Watson (DW) test for first order correlation and the Q and Lagrange Multiplier (LM) tests for remaining ARCH effects.

3.3.1 GenBM Data

We will now present the results obtained when fitting the various models of Section 3.2 to the GenBM data. Since we are fitting the models to data generated using known parameter values, this section, together with Section 3.3.2, serves as a test of our programs as well as an illustration of the underlying theory. Of course the normal and BM models should fit well since the data was actually generated by these models. The NIG and BIG models should also fit well since they nest the normal and BM models. We can also inspect changes in accuracy of the parameter estimates as more and more data is used.

ML estimates

Table 3.2 gives the ML estimates of the parameters, with asymptotic standard errors in brackets, of the D-NOR, D-NIG, HLC-BM, HLC-BIG, ID-BM and ID-BIG model fits to the GenBM data (since the D-NOR, HLC-BIG and ID-BIG models assume $\psi \to \infty$, these models provide no estimate for ψ). Indeed, we see that all the models fit well. Taking standard errors into account (and ignoring ψ for now), we see that most of the estimates are indeed within one standard error of the true values, with a few estimates being only slightly more than one standard error away from the true values.

Note that we have bounded the estimates of ψ above by 25 since larger values cause computational problems, and that the D-NIG estimate for ψ is at this boundary. However, when $\psi = 25$ the $SNIG(\beta,\psi)$-distribution is virtually identical to the $N(\beta,1)$-distribution (consider Figure 2.3). In other words, for all practical purposes, a value of $\psi = 25$ is as good as $\psi \to \infty$. The same holds for the ID-BIG5 model estimate of $\psi \approx 22.3444$. It seems that the HLC-BIG model (and to a lesser extent the ID-BIG30 model) significantly underestimates ψ. We found that this seems to be the case in general for data generated with high values of ψ. However, when we consider Figure 2.3 we see that even for relatively small values of ψ (such as $\psi \approx 3.6984$ estimated by the HLC-BIG model) we would not expect the $SNIG(\beta,\psi)$-distribution to differ much from the $N(\beta,1)$-distribution. Indeed, Figure 3.1 shows that is true for the case when $\beta \approx 0.1481$, as estimated by the HLC-BIG model. Thus, the fact that ψ is underestimated by the HLC-BIG model is not a critical point.

Note that, although all the models fit the data well, the models based on NIG-distributed innovations (or underlying BIG processes) generally lead to slightly smaller standard errors. Furthermore, the standard errors of the estimates obtained by the D-NOR and D-NIG fits are considerably larger than those obtained by the HLC-BM and HLC-BIG fits which in turn are larger than the standard errors obtained from the ID-BM and ID-BIG fits. In addition, the standard errors resulting from the ID-BM and ID-BIG model fits utilising 5 minute data are noticeably smaller than those resulting from the models utilising 30 minute data. This illustrates the accuracy gain as more data is used.

Quality of fit

We now check the quality of each fit, starting with the D-NOR and D-NIG models which utilise only daily return data, moving on to the HLC-BM and HLC-BIG models which also utilise daily high and low return and concluding with the ID-BM and ID-BIG models which utilise intraday day.

From Figure 3.2 we see that the distributions of the $\hat{X}_n$ and $\hat{Z}_n$ residuals from the D-NOR fit to the GenBM data correspond very well with their model implied distributions. Thus, the D-NOR model fits the GenBM data very well, as expected. Figure 3.3 shows that the D-NIG model also fits the GenBM data very well, since the $\hat{X}_n$ residuals and the $\hat{Z}_{X,n}$ and $\hat{W}_{X,n}$ estimates correspond to their model implied distributions very well.

Parameter	GenBM	D-NOR	D-NIG
β	0.20	0.1159 (0.3579)	0.1151 (0.1320)
ψ	∞	-	25.0000 (3.8344)
ν	0.00	0.0014 (0.0038)	0.0014 (0.0013)
ϕ	0.15	0.1291 (0.0321)	0.1288 (0.0318)
$\alpha_0 x 10^5$	1.00	0.7968 (0.3028)	0.8106 (0.2968)
α_1	0.10	0.1030 (0.0216)	0.1036 (0.0213)
β_1	0.80	0.8281 (0.0391)	0.8264 (0.0379)
Parameter	**GenBM**	**HLC-BM**	**HLC-BIG**
β	0.20	0.1742 (0.0637)	0.1481 (0.0535)
ψ	∞	-	3.6984 (0.1540)
ν	0.00	0.0007 (0.0005)	0.0010 (0.0005)
ϕ	0.15	0.1574 (0.0272)	0.1546 (0.0266)
$\alpha_0 x 10^5$	1.00	0.8889 (0.1131)	0.8591 (0.0846)
α_1	0.10	0.0988 (0.0086)	0.1040 (0.0090)
β_1	0.80	0.8076 (0.0165)	0.8115 (0.0141)
Parameter	**GenBM**	**ID-BM30**	**ID-BIG30**
β	0.20	0.2163 (0.0556)	0.2138 (0.0552)
ψ	∞	-	6.8123 (1.2553)
ν	0.00	0.0004 (0.0004)	0.0004 (0.0004)
ϕ	0.15	0.1294 (0.0264)	0.1294 (0.0265)
$\alpha_0 x 10^5$	1.00	0.8815 (0.1057)	0.8857 (0.1050)
α_1	0.10	0.0925 (0.0072)	0.0925 (0.0072)
β_1	0.80	0.8098 (0.0155)	0.8093 (0.0154)
Parameter	**GenBM**	**ID-BM5**	**ID-BIG5**
β	0.20	0.2348 (0.0358)	0.2354 (0.0360)
ψ	∞	-	22.3444 (3.0278)
ν	0.00	0.0001 (0.0002)	0.0001 (0.0002)
ϕ	0.15	0.1407 (0.0148)	0.1407 (0.0148)
$\alpha_0 x 10^5$	1.00	0.9936 (0.0427)	0.9928 (0.0425)
α_1	0.10	0.1024 (0.0032)	0.1025 (0.0032)
β_1	0.80	0.7972 (0.0061)	0.7971 (0.0061)

Table 3.2: ML estimates of all model fits to GenBM data.

Turning to the HLC-BM model, we see from Figure 3.4 that the distribution of the $\hat{X}_n$ and $\hat{Z}_n$ residuals agree very closely with their model implied normal distributions, so that the HLC-BM model fits the GenBM data very well, as expected. From Figures 3.5 and 3.6 we see that the distributions of $\hat{X}_n$, $\hat{Z}_{X,n}$ and $\hat{W}_{X,n}$, obtained from the HLC-BIG fit to the GenBM data, also agree quite well with their model implied distributions, while the distribution of $\hat{W}_{XAB,n}$ corresponds to its model implied distribution slightly less well. Remember that we have assumed that $\mu_n = 0$ in order

to calculate the estimates $\hat{W}_{XAB,n}$. In this case the true range of the μ_n's is $-0.0043 < \mu_n < 0.0058$ with an average value of $\mu_n = 0.0006$. Therefore, the $\mu_n = 0$ assumption is not the main reason why the distribution of $\hat{W}_{XAB,n}$ differs from its model implied distribution in this case. Rather, the relatively small estimate for the parameter $\hat{\psi} \approx 3.6984$ causes the $\hat{W}_{XAB,n}$ values to be more spread out around 1 than they theoretically should be (although, as argued above, the difference is not critical).

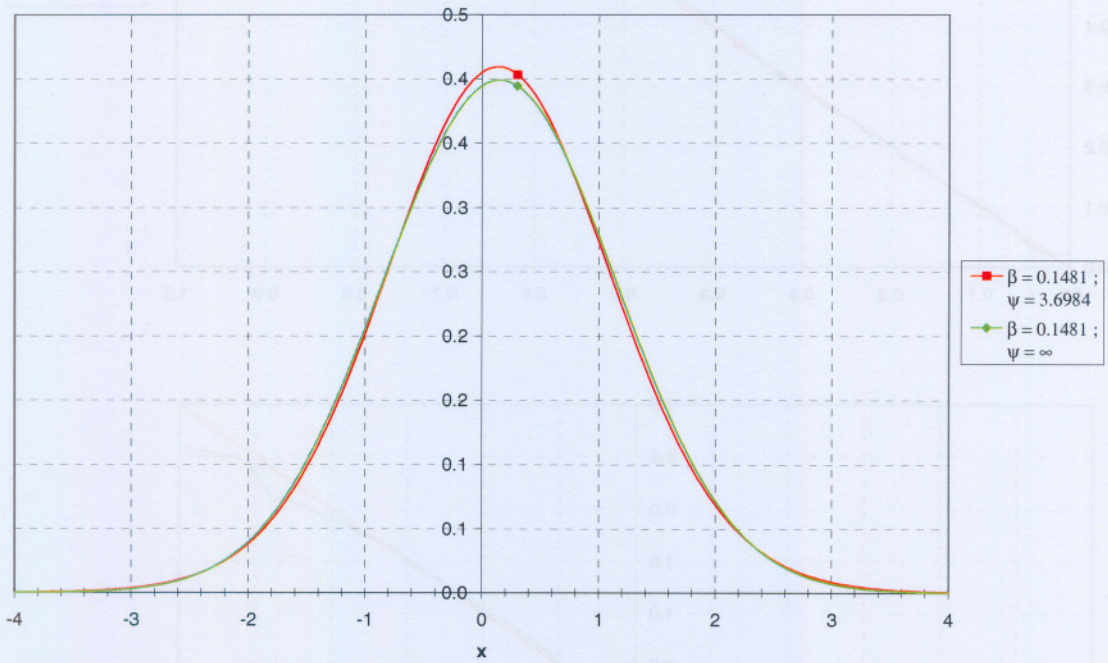

Figure 3.1: $SNIG(0.1481, \psi)$-densities.

Figures 3.7 and 3.8 show the PIT and Q-Q plots for $\hat{X}_n$, $\hat{Z}_n$, $\hat{V}_n$ and $\hat{T}_n$, checking the fits of the ID-BM model to the 5 and 30 minute GenBM data. We see that the models fit the data reasonably well. Note that, as expected, the model fits better with the 5 minute data than with the 30 minute data. Next, we turn to the ID-BIG model. Figures 3.9 and 3.10 show that $\hat{X}_n$, $\hat{V}_n$, $\hat{T}_n$, $\hat{W}_{XV,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ all correspond very well with their model-implied distributions. Thus, as expected, the ID-BIG model fits the 5 minute GenBM data very well. The corresponding figures for the ID-BIG fit to the 30 minute GenBM data are very similar to their 5 minute counterparts. The PIT and Q-Q plots in Figures 3.11 and 3.12 show that the ID-BIG model also fits the 30 minute data very well, although not quite as well as the 5 minute data.

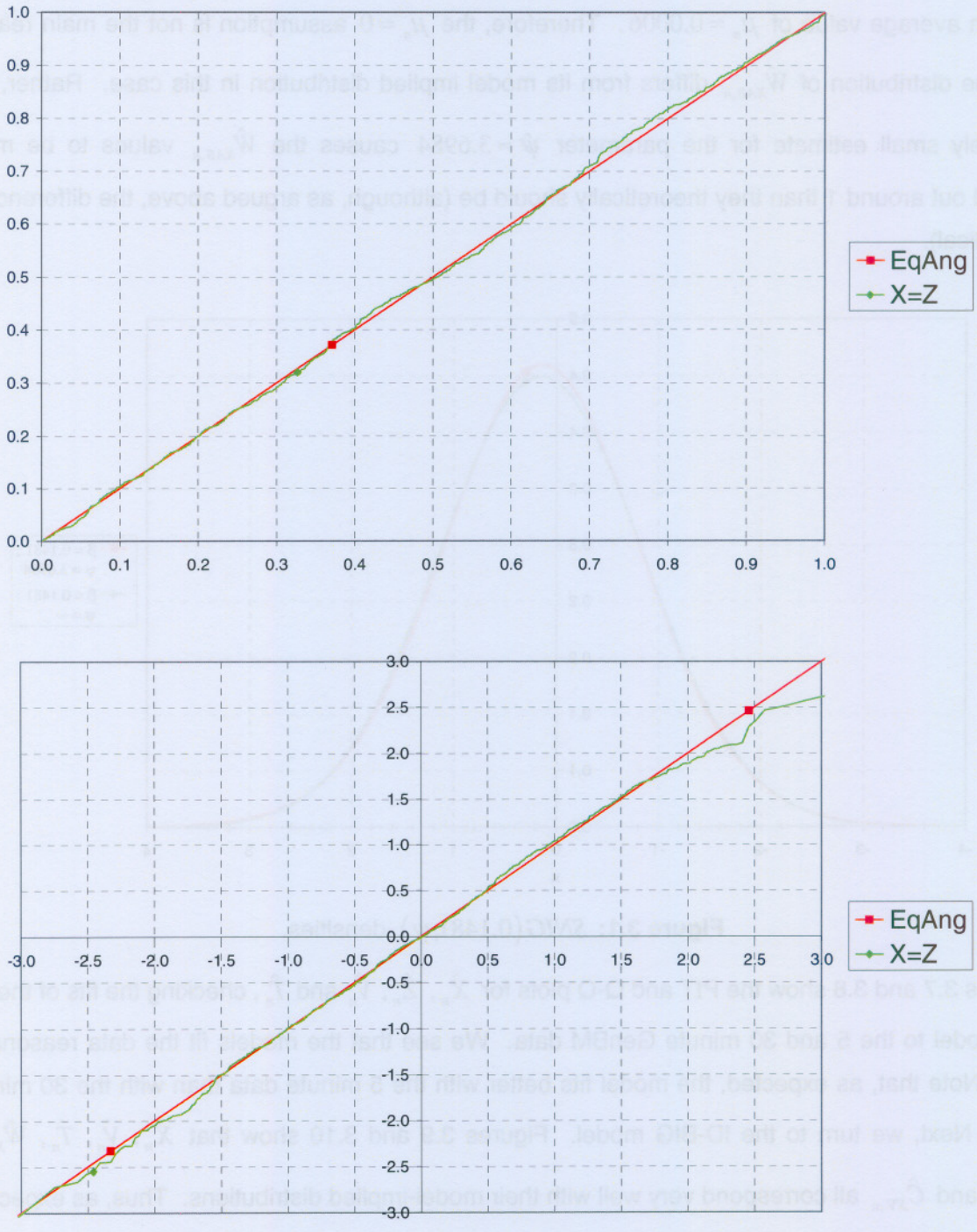

Figure 3.2: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$ and $\hat{Z}_n$ checking D-NOR fit to GenBM data.

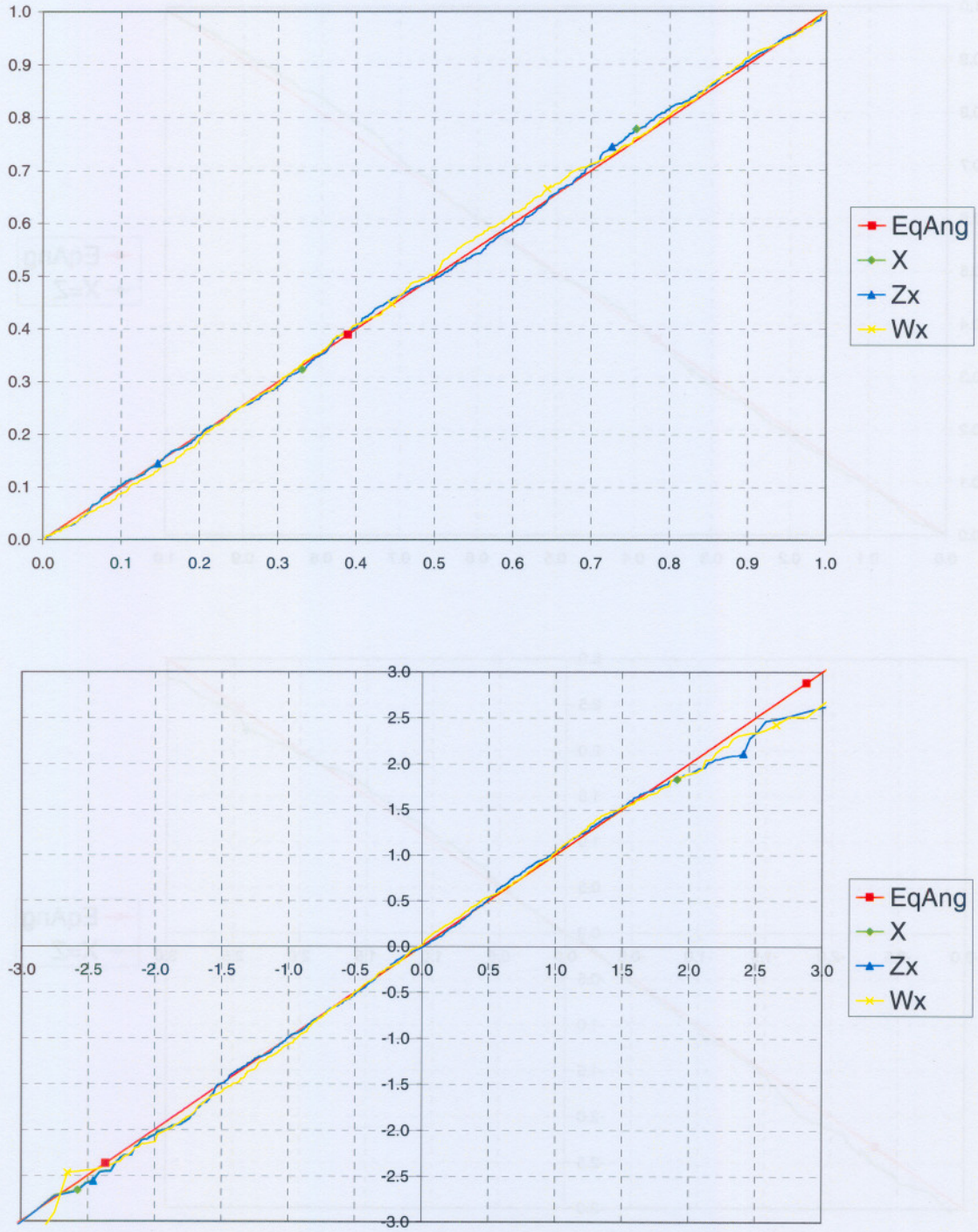

Figure 3.3: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_{X,n}$ and $\hat{W}_{X,n}$ checking D-NIG fit to GenBM data.

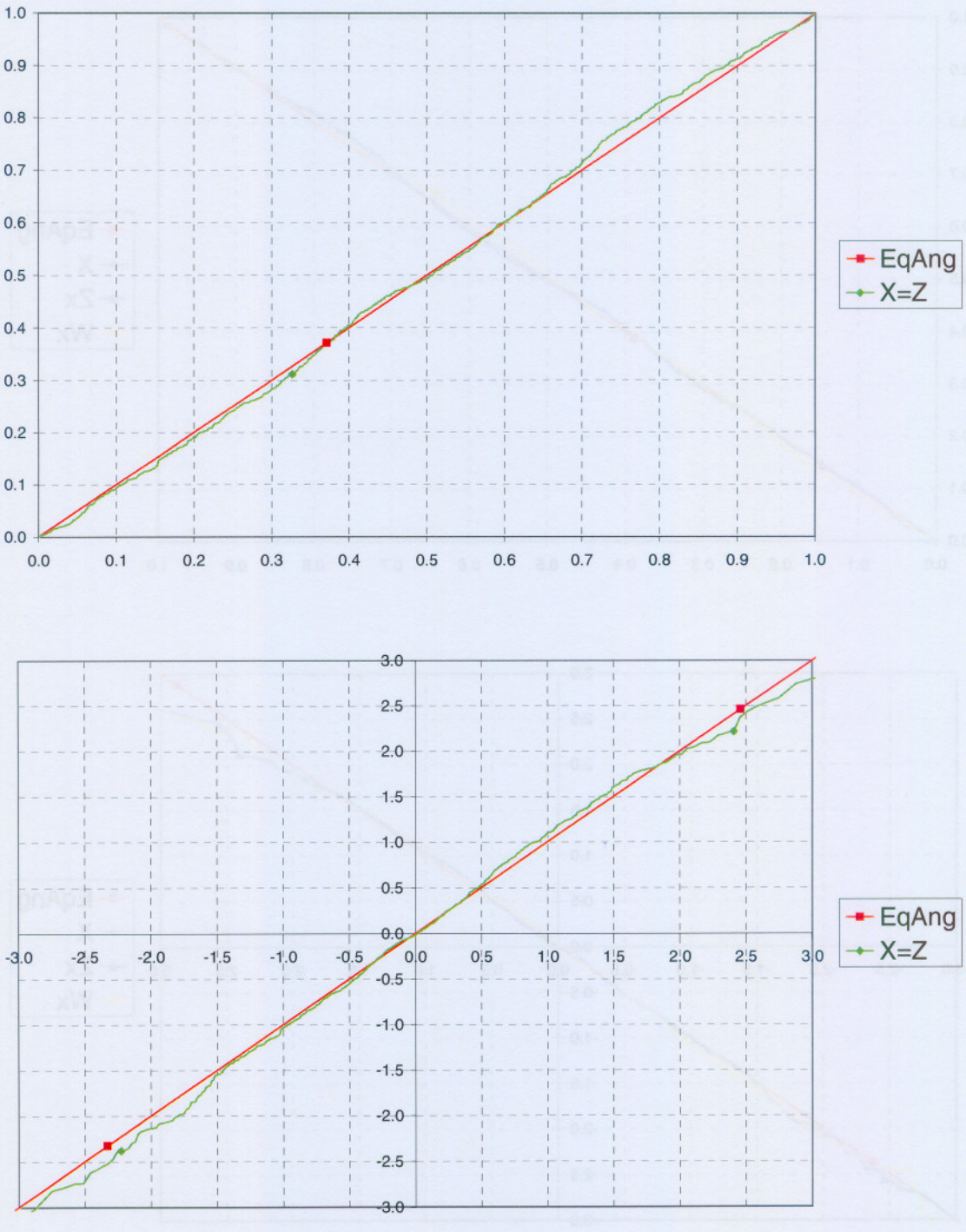

Figure 3.4: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$ and $\hat{Z}_n$ checking HLC-BM fit to GenBM data.

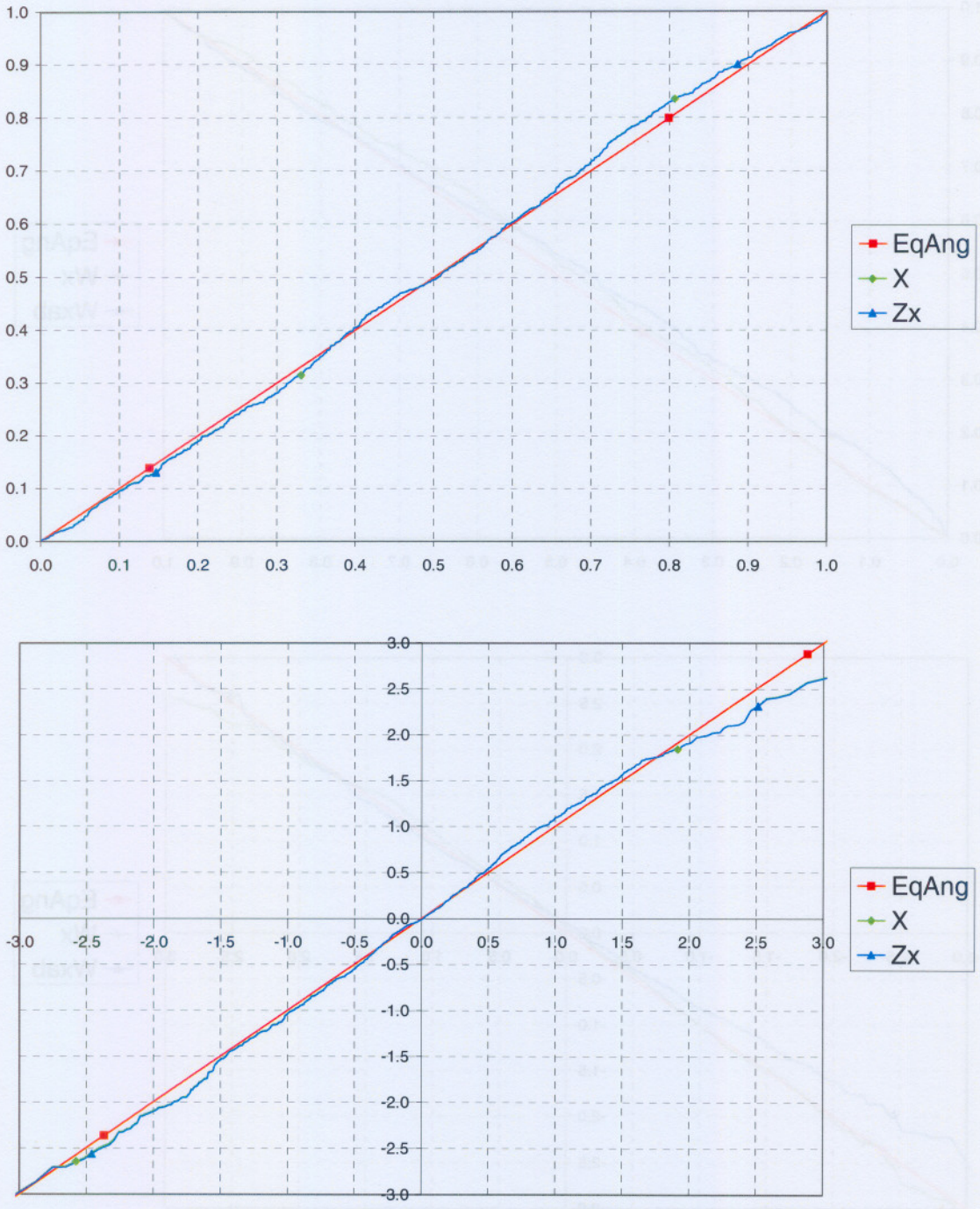

Figure 3.5: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$ and $\hat{Z}_{X,n}$ checking HLC-BIG fit to GenBM data.

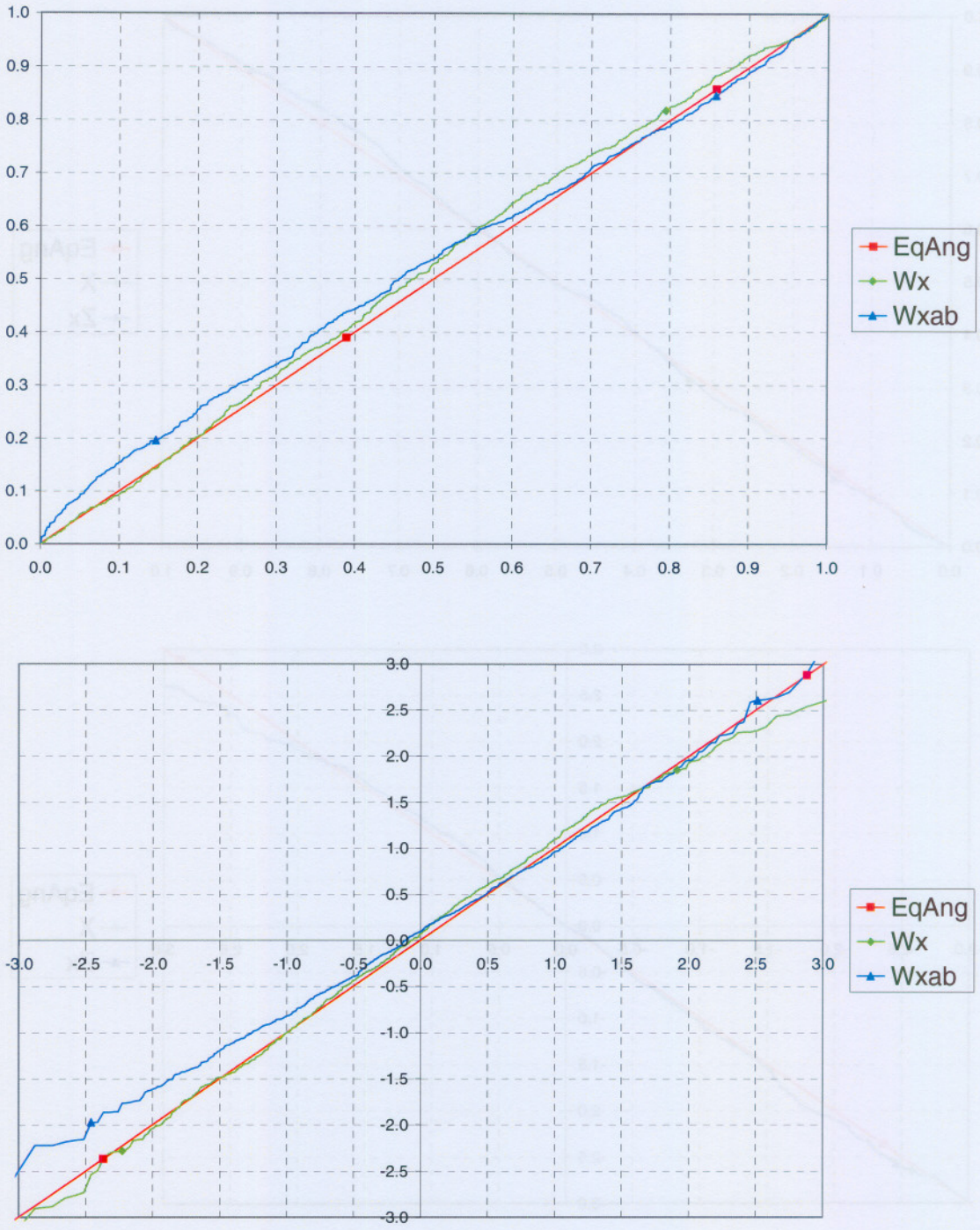

Figure 3.6: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{X,n}$ and $\hat{W}_{XAB,n}$ checking HLC-BIG fit to GenBM data.

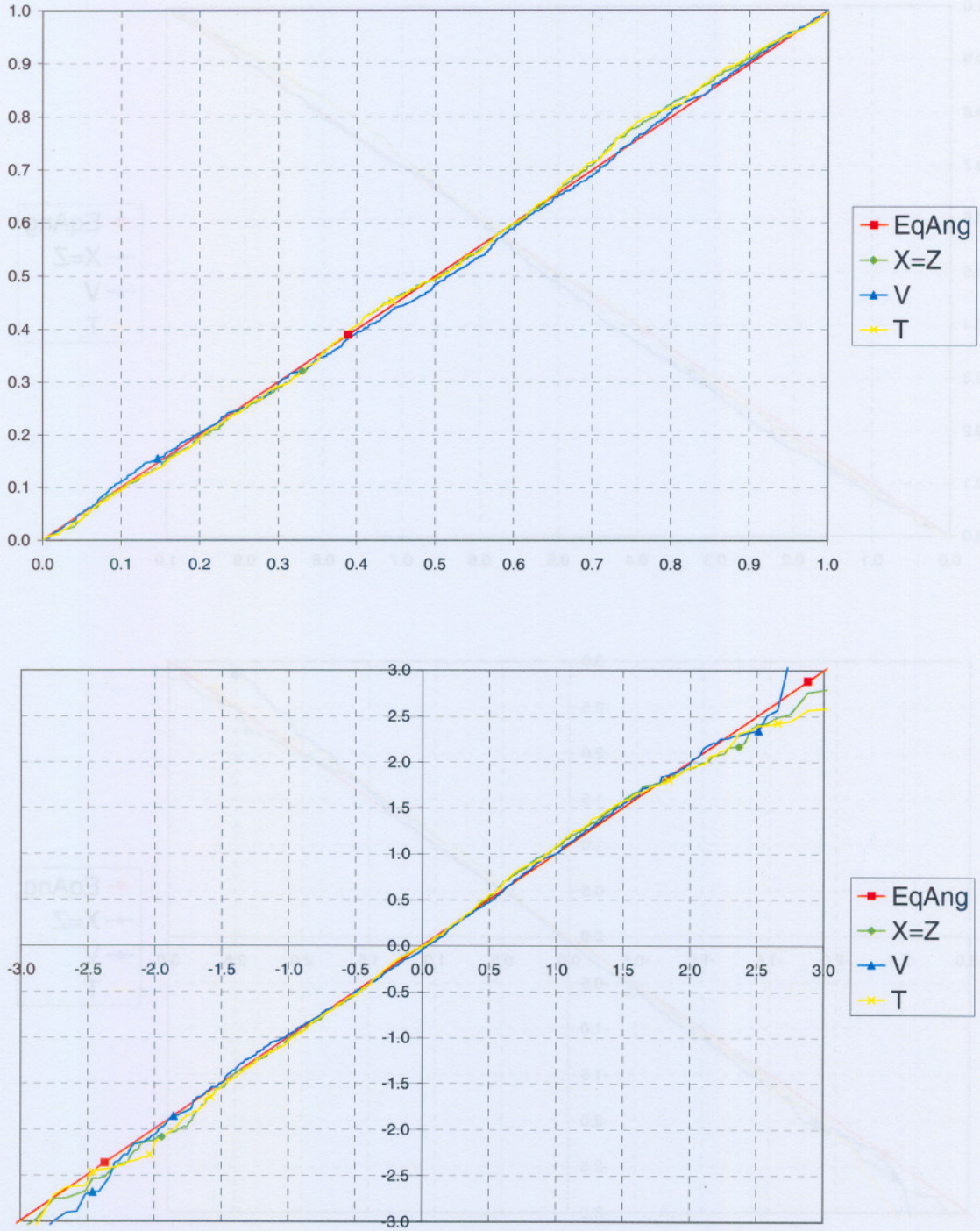

Figure 3.7: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BM5 fit to GenBM data.

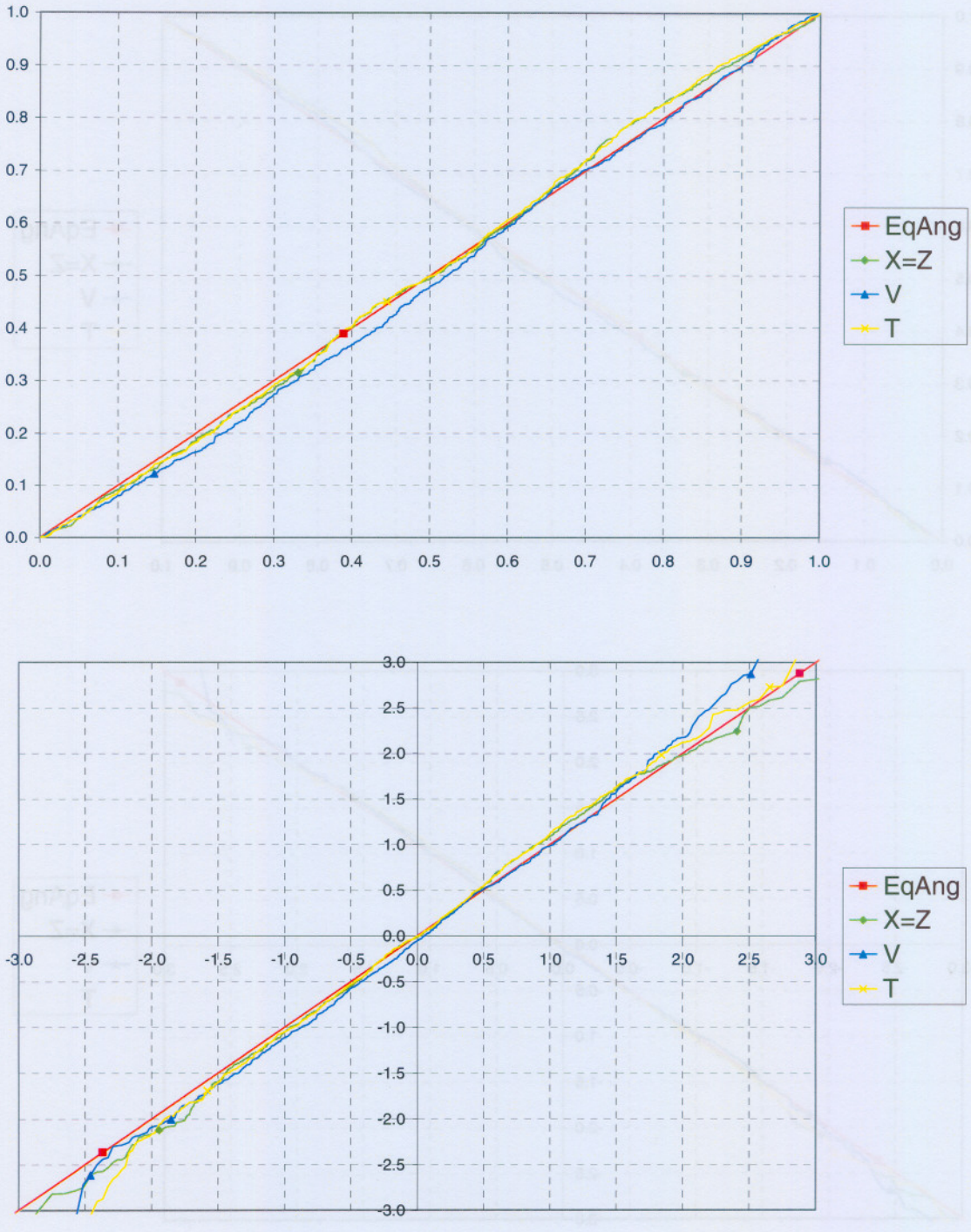

Figure 3.8: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BM30 fit to GenBM data.

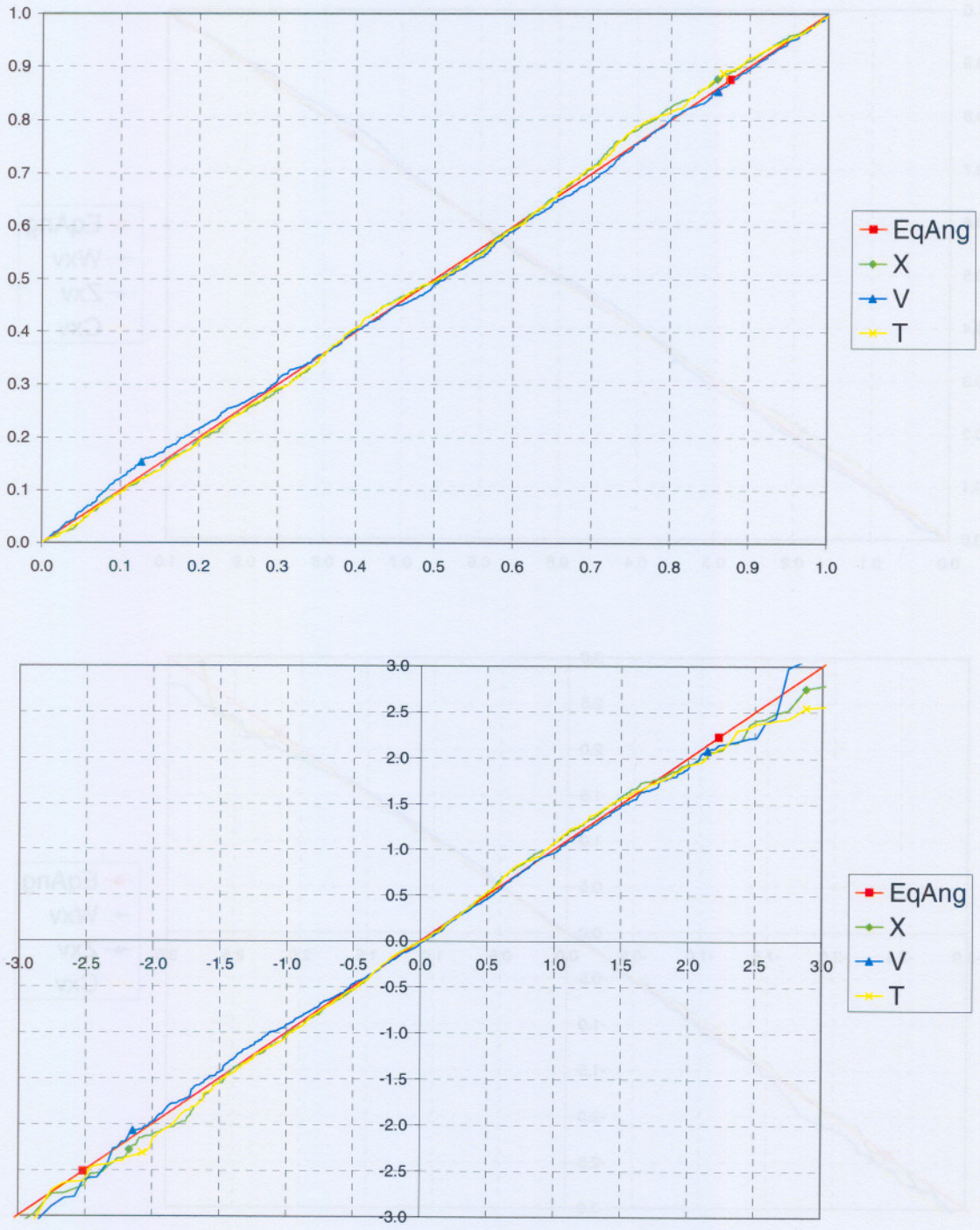

Figure 3.9: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BIG5 fit to GenBM data.

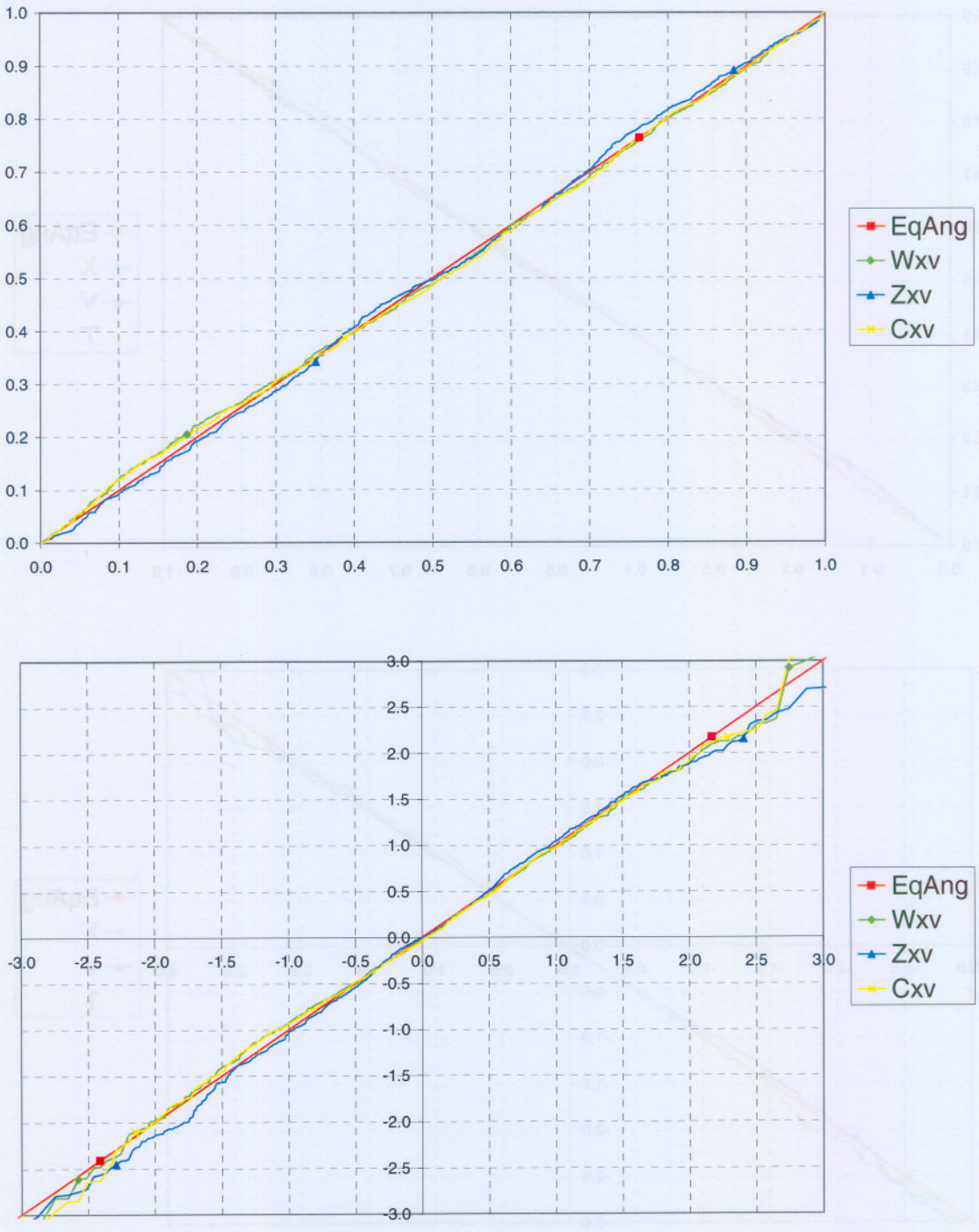

Figure 3.10: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{XV,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking ID-BIG5 fit to GenBM data.

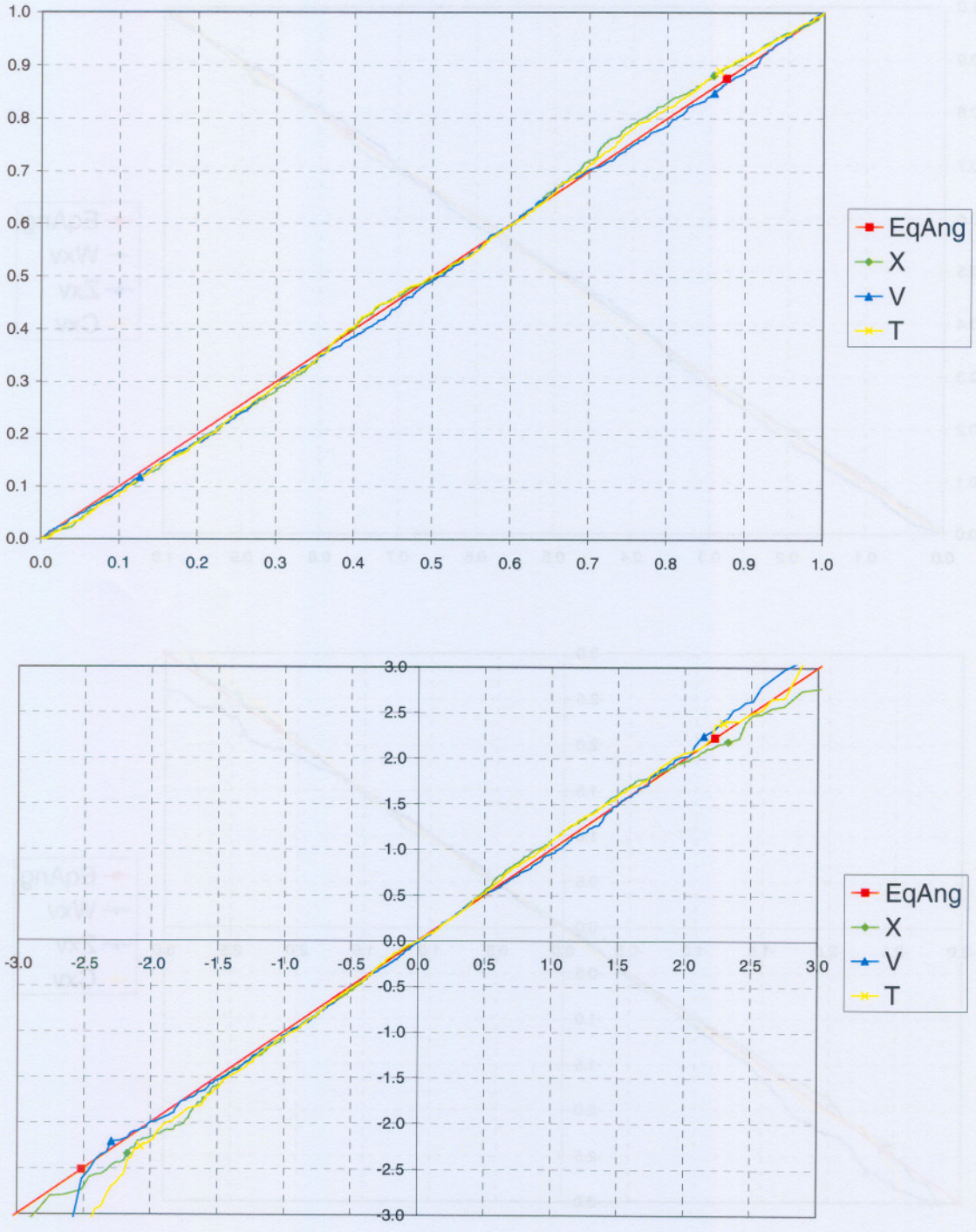

Figure 3.11: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BIG30 fit to GenBM data.

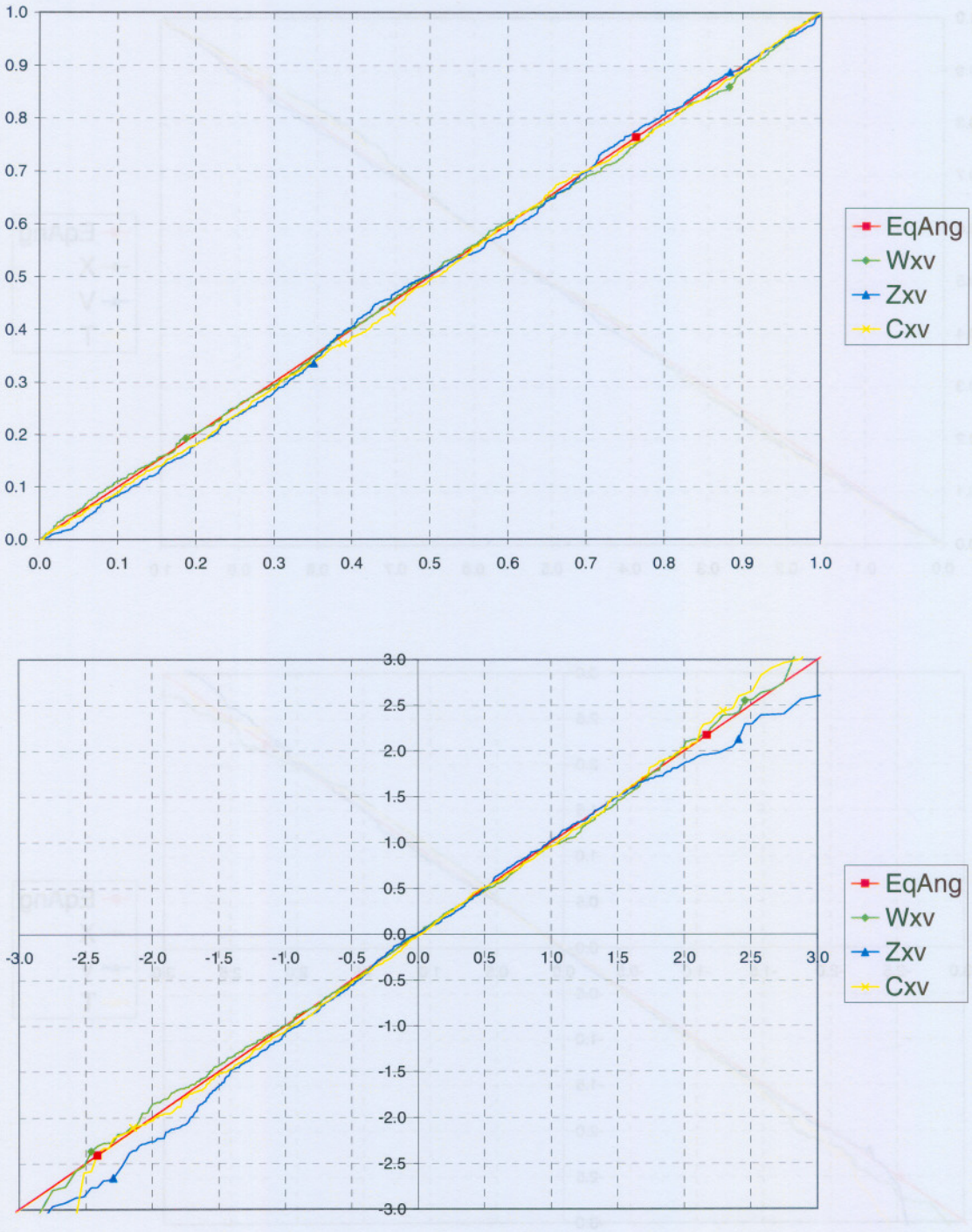

Figure 3.12: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{XV,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking ID-BIG30 fit to GenBM data.

Volatility estimates

Next, we examine the volatility estimates obtained from the D-NIG, HLC-BIG and ID-BIG model fits to the GenBM data. We see from the top panel of Figure 3.13 (only the values for days 500 to 600 are displayed to make the details visible on the graph) that the $\hat{W}_{X,n}$ estimates differ only very slightly from the value of 1, which in effect was used to generate the GenBM data. The lower boundary of $\hat{W}_{X,n}$, as described in Section 2.3.5, is also clear. In this case this lower boundary actually means that $\hat{W}_{X,n}$ cannot be much lower than the theoretical value of 1, improving the accuracy of $\hat{W}_{X,n}$. Since the parameter estimate $\hat{\psi} \approx 25$ is quite large we would indeed expect the values of $\hat{W}_{X,n}$ to be very close to 1 (see Figure 2.1). Thus, in this case the ordinary GARCH volatility estimator $\hat{h}_n$ is hardly adjusted at all by the adjusted volatility estimator $\hat{h}_n\hat{W}_{X,n}$, as is seen from the bottom panel of Figure 3.13 where the graphs of $\hat{h}_n$ and $\hat{h}_n\hat{W}_{X,n}$ are so similar as to be indistinguishable. Although the realised volatility R_n is not incorporated into the D-NIG model, we show it here to illustrate the difference between R_n, $\hat{h}_n$ and $\hat{h}_n\hat{W}_{X,n}$ as estimators of the same true but unknown underlying volatility, and also in order to compare the GenBM results with the GenBIG results in the next section.

Figure 3.14 displays the volatility estimates obtained from the HLC-BIG fit to the GenBM data. The lower bound of $\hat{W}_{X,n}$, as well as the spread of $\hat{W}_{X,n}$ and $\hat{W}_{XAB,n}$ around 1 can be seen in the top panel of Figure 3.14. Again, the lower bound on $\hat{W}_{X,n}$ actually increases its accuracy in the case of data generated using normally distributed innovations. The bottom panel of Figure 3.14 shows the difference between $\hat{h}_n\hat{W}_{X,n}$ ($\hat{h}_n$ alone is not shown, since it is very close to $\hat{h}_n\hat{W}_{X,n}$), $\hat{h}_n\hat{W}_{XAB,n}$ and R_n as estimators for the true underlying volatility.

Comparing these graphs with the volatility estimates obtained from the ID-BIG fit to the 5 minute data, we see from the top panel of Figure 3.15 that the values of $\hat{W}_{XV,n}$ are not as far away from the theoretical value of 1 as was the case with the values of $\hat{W}_{X,n}$ and $\hat{W}_{XAB,n}$ obtained from the HLC-BIG fit. The bottom panel of Figure 3.15 shows that $\hat{W}_{XV,n}$, being so close to 1, do not notably adjust the ordinary GARCH volatility estimator $\hat{h}_n$. This effect, or lack thereof, is also shown in Figure 3.16, where it is clear that $\hat{h}_n$ and $\hat{h}_n\hat{W}_{XV,n}$ are very nearly identical estimates of

the underlying volatility. Comparing the spread of $\hat{h}_n$ around the equi-angular line we see that, although there is some difference between $\hat{h}_n$ and the realised volatility R_n, they are reasonably close to each other. In fact, the correlation between $\hat{h}_n$ and R_n is 0.9024, which is reasonably high. This illustrates the point that the ordinary GARCH volatility estimator $\hat{h}_n$ is a smoothed version of the more highly variable daily actual volatility (see also the volatility estimate discussions in Sections 3.3.2 and 3.3.3).

As for the volatility estimates obtained from the ID-BIG model fit to the 30 minute GenBM data, Figure 3.18 shows that the adjustment that $\hat{W}_{XV,n}$ makes on $\hat{h}_n$ is again not prominent in this case, but both $\hat{h}_n$ and $\hat{h}_n\hat{W}_{XV,n}$ differ much more from R_n (on the equi-angular line) than in the 5 minute data case. This illustrates the reasoning just below (3.2.18) where we stated that $R_n \approx \hat{h}_n\hat{W}_{XV,n}$ if I is large. Furthermore, in this case the $\hat{W}_{XV,n}$'s are very close to 1, so that R_n should be close to $\hat{h}_n$ for large I. $I=78$ in the 5 minute case while $I=13$ in the 30 minute case, a considerable difference which is the source of the difference between the spreads of $\hat{h}_n$ around R_n in the 5 minute and 30 minute cases.

Testing for independence

Tables 3.3 and 3.4 contain the results of the DW, Q and LM tests on all the residuals and estimates of the model fits to the GenBM data considered in this section. Since $W_{X,n}$, $W_{XAB,n}$, $W_{XV,n}$, V_n and $C_{XV,n}$ are positive, and for reasons explained at the end of Section A.2 in the Appendix, we perform the tests on the natural logarithm of these residuals and estimates. In all cases the results indicate no first order correlation or remaining ARCH effects.

All in all, it seems that the models assuming normally distributed innovations (or underlying BM processes), as well as those assuming NIG-distributed innovations (or underlying BIG processes) fit the GenBM data quite well, as expected.

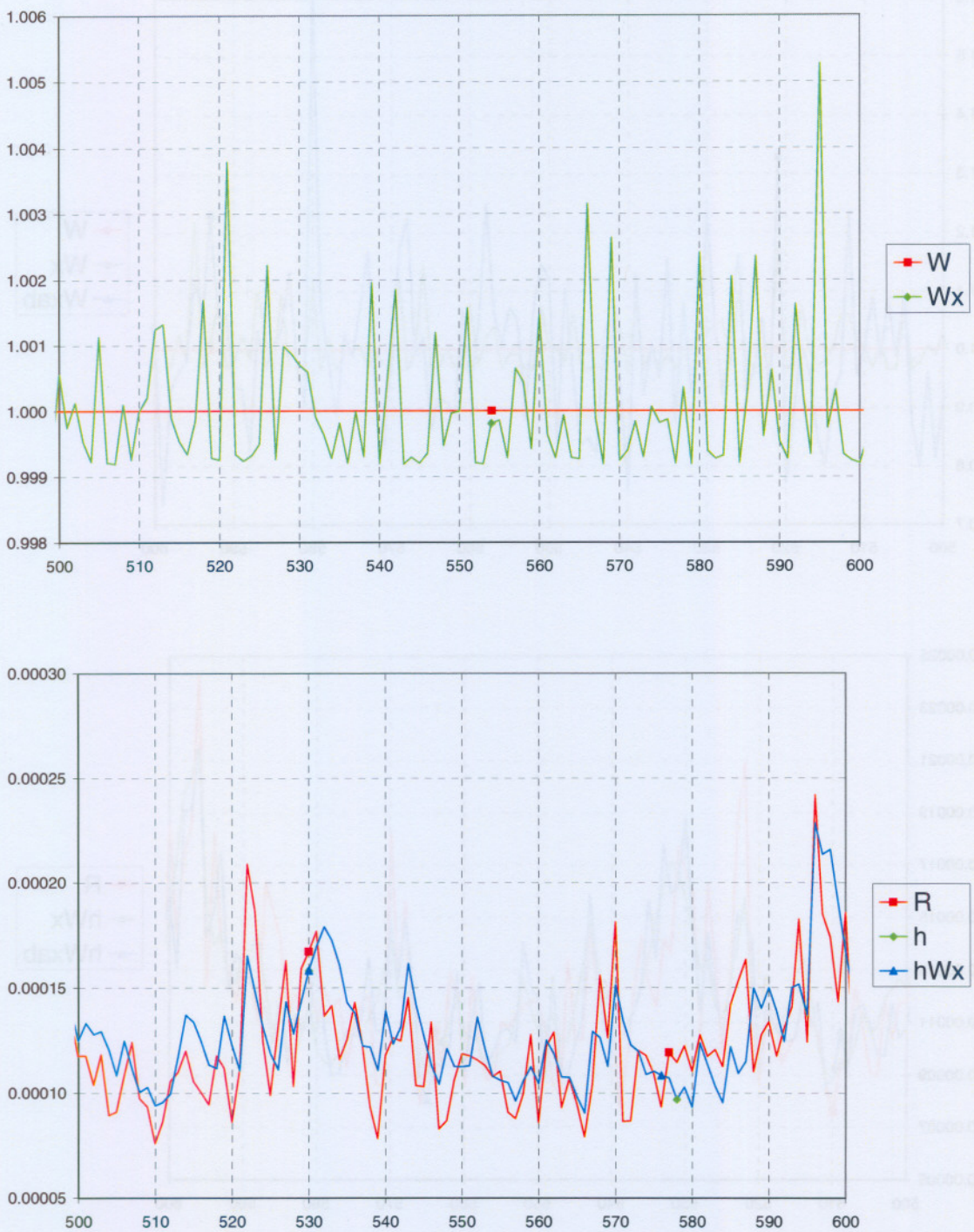

Figure 3.13: Plots of W_n and $\hat{W}_{X,n}$ (top panel) and R_n, $\hat{h}_n$, and $\hat{h}_n\hat{W}_{X,n}$ (bottom panel) for D-NIG fit to GenBM data.

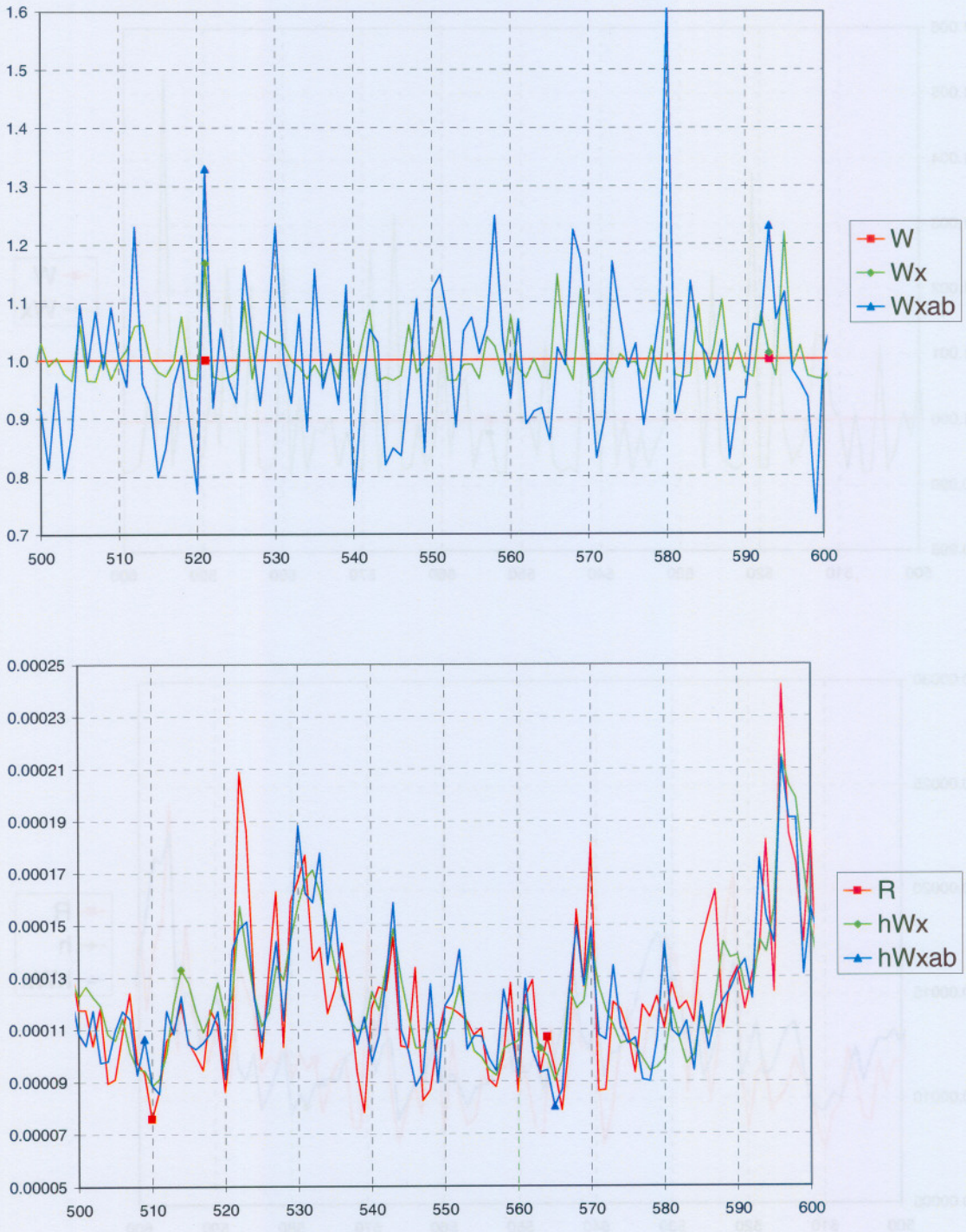

Figure 3.14: Plots of W_n, $\hat{W}_{X,n}$ and $\hat{W}_{XAB,n}$ (top panel) and R_n, $\hat{h}_n\hat{W}_{X,n}$ and $\hat{h}_n\hat{W}_{XAB,n}$ (bottom panel) for HLC-BIG fit to GenBM data.

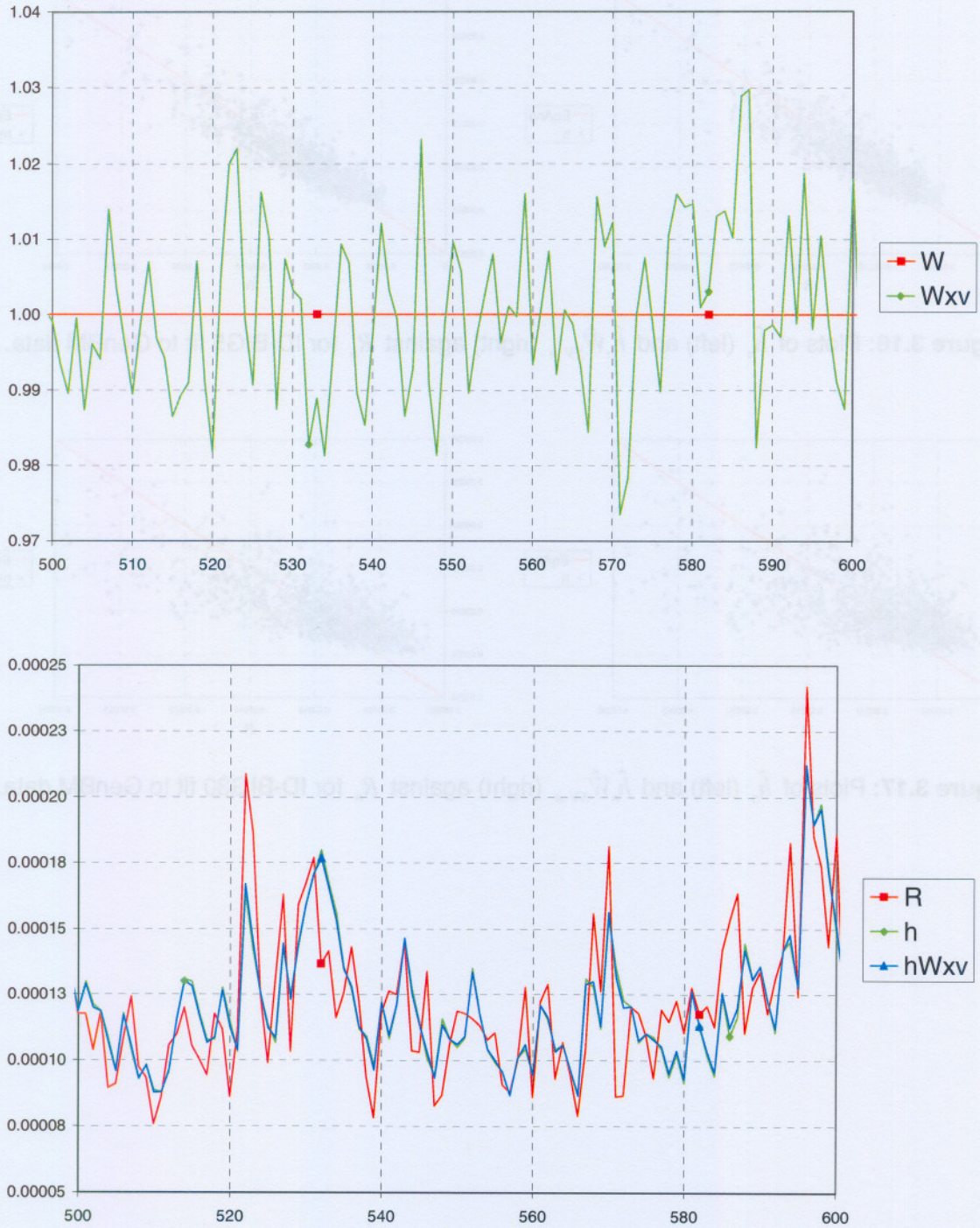

Figure 3.15: Plots of W_n and $\hat{W}_{XV,n}$ (top panel) and R_n, $\hat{h}_n$, and $\hat{h}_n\hat{W}_{XV,n}$ (bottom panel) for ID-BIG5 fit to GenBM data.

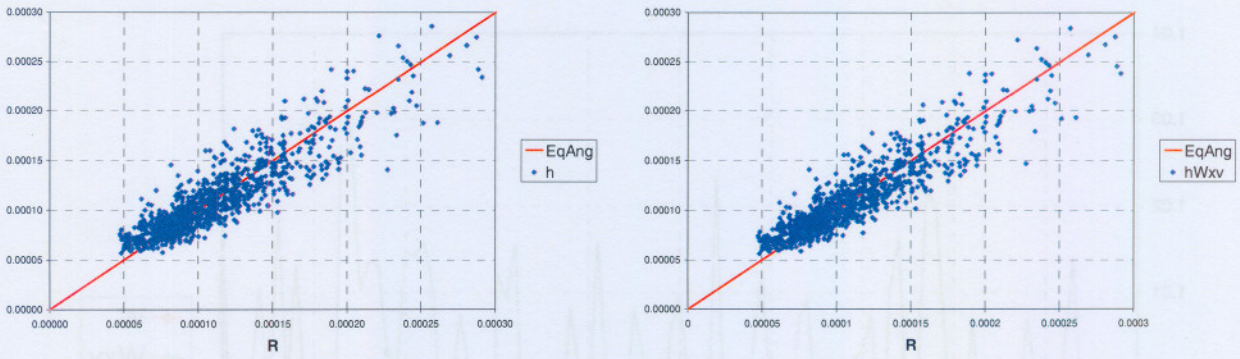

Figure 3.16: Plots of $\hat{h}_n$ (left) and $\hat{h}_n \hat{W}_{XV,n}$ (right) against R_n for ID-BIG5 fit to GenBM data.

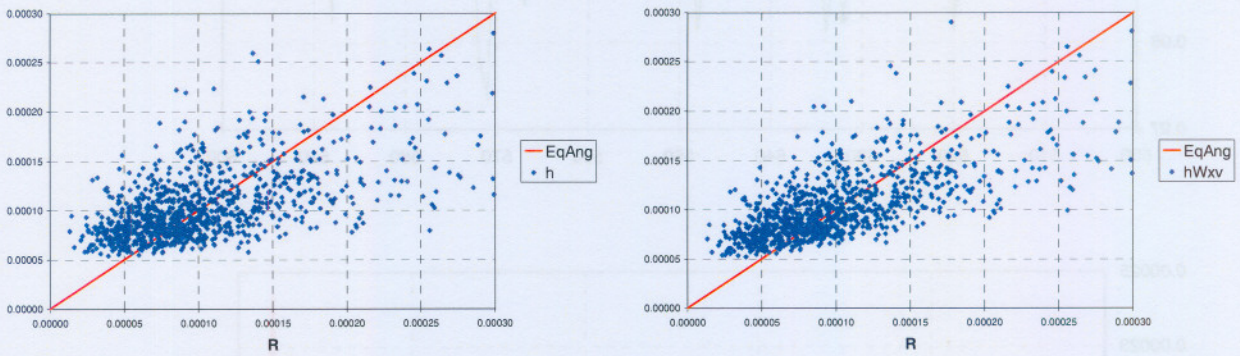

Figure 3.17: Plots of $\hat{h}_n$ (left) and $\hat{h}_n \hat{W}_{XV,n}$ (right) against R_n for ID-BIG30 fit to GenBM data.

D-NOR	X=Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.018	0.893	0.014	0.905
2	0.313	0.855	0.303	0.859
3	0.317	0.957	0.308	0.959
4	1.821	0.769	1.843	0.765
5	2.554	0.768	2.551	0.769
6	3.200	0.783	3.271	0.774
7	3.660	0.818	3.737	0.810
8	5.350	0.720	5.372	0.717
9	5.368	0.801	5.378	0.800
10	7.114	0.715	6.816	0.743
Durbin-Watson				2.007
Pr<DW				0.547
Pr>DW				0.453

D-NIG	Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.015	0.901	0.012	0.914
2	0.262	0.878	0.253	0.881
3	0.277	0.964	0.267	0.966
4	1.901	0.754	1.928	0.749
5	2.709	0.745	2.710	0.745
6	3.360	0.763	3.431	0.753
7	3.838	0.798	3.910	0.790
8	5.640	0.688	5.662	0.685
9	5.660	0.773	5.668	0.773
10	7.539	0.674	7.209	0.706
Durbin-Watson				2.005
Pr<DW				0.534
Pr>DW				0.466

D-NIG	X			
Order	Q	Pr>Q	LM	Pr>LM
1	0.008	0.928	0.006	0.940
2	0.261	0.878	0.254	0.881
3	0.269	0.966	0.261	0.967
4	1.811	0.771	1.837	0.766
5	2.528	0.772	2.534	0.771
6	3.178	0.786	3.258	0.776
7	3.629	0.821	3.713	0.812
8	5.328	0.722	5.370	0.717
9	5.346	0.803	5.375	0.801
10	7.065	0.719	6.794	0.745
Durbin-Watson				2.006
Pr<DW				0.539
Pr>DW				0.461

D-NIG	log Wx			
Order	Q	Pr>Q	LM	Pr>LM
1	0.609	0.435	0.604	0.437
2	0.665	0.717	0.655	0.721
3	0.825	0.844	0.798	0.850
4	1.441	0.837	1.394	0.845
5	10.209	0.070	10.395	0.065
6	10.335	0.111	10.462	0.107
7	10.491	0.162	10.564	0.159
8	10.616	0.224	10.661	0.222
9	11.570	0.239	12.084	0.209
10	12.926	0.228	12.786	0.236
Durbin-Watson				2.023
Pr<DW				0.642
Pr>DW				0.358

HLC-BM	X=Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.187	0.665	0.174	0.677
2	0.975	0.614	0.958	0.619
3	0.978	0.807	0.964	0.810
4	1.990	0.738	1.973	0.741
5	3.117	0.682	3.038	0.694
6	3.409	0.756	3.377	0.760
7	3.759	0.807	3.725	0.811
8	5.654	0.686	5.486	0.705
9	5.692	0.770	5.505	0.788
10	7.290	0.698	6.791	0.745
Durbin-Watson				2.044
Pr<DW				0.759
Pr>DW				0.242

HLC-BIG	Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.423	0.516	0.399	0.528
2	1.381	0.501	1.368	0.505
3	1.381	0.710	1.372	0.712
4	2.515	0.642	2.493	0.646
5	3.991	0.551	3.860	0.570
6	4.269	0.640	4.171	0.654
7	4.675	0.700	4.568	0.713
8	6.814	0.557	6.521	0.589
9	6.851	0.653	6.541	0.685
10	8.883	0.543	8.132	0.616
Durbin-Watson				2.039
Pr<DW				0.732
Pr>DW				0.268

HLC-BIG	X			
Order	Q	Pr>Q	LM	Pr>LM
1	0.325	0.569	0.307	0.580
2	1.316	0.518	1.302	0.521
3	1.328	0.723	1.322	0.724
4	2.255	0.689	2.230	0.694
5	3.479	0.627	3.381	0.642
6	3.753	0.710	3.693	0.718
7	4.119	0.766	4.057	0.773
8	6.017	0.645	5.798	0.670
9	6.057	0.734	5.819	0.758
10	7.712	0.657	7.138	0.712
Durbin-Watson				2.043
Pr<DW				0.749
Pr>DW				0.251

HLC-BIG	log Wx			
Order	Q	Pr>Q	LM	Pr>LM
1	0.799	0.372	0.793	0.373
2	0.848	0.654	0.838	0.658
3	1.000	0.801	0.969	0.809
4	1.766	0.779	1.714	0.788
5	6.269	0.281	6.433	0.266
6	6.347	0.386	6.481	0.372
7	6.650	0.466	6.717	0.459
8	6.808	0.557	6.861	0.552
9	7.978	0.536	8.489	0.486
10	10.181	0.425	10.131	0.429
Durbin-Watson				2.025
Pr<DW				0.655
Pr>DW				0.345

HLC-BIG	log Wxab			
Order	Q	Pr>Q	LM	Pr>LM
1	0.325	0.569	0.319	0.572
2	1.273	0.529	1.252	0.535
3	1.830	0.609	1.821	0.610
4	1.830	0.767	1.823	0.768
5	2.253	0.813	2.253	0.813
6	2.275	0.893	2.292	0.891
7	2.276	0.943	2.292	0.942
8	5.319	0.723	5.588	0.693
9	8.837	0.452	9.053	0.432
10	9.019	0.530	9.300	0.504
Durbin-Watson				2.021
Pr<DW				0.629
Pr>DW				0.371

Table 3.3: DW, Q and LM tests for daily and HLC model fits to GenBM data.

ID-BM5	X=Z				ID-BM5	log V				ID-BM5	T			
Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM
1	0.123	0.725	0.113	0.737	1	0.027	0.870	0.022	0.881	1	0.009	0.925	0.006	0.937
2	0.802	0.670	0.786	0.675	2	4.559	0.102	4.576	0.102	2	0.548	0.760	0.534	0.766
3	0.805	0.848	0.790	0.852	3	5.044	0.169	5.097	0.165	3	0.726	0.867	0.711	0.871
4	1.930	0.749	1.921	0.750	4	5.564	0.234	5.781	0.216	4	2.546	0.636	2.549	0.636
5	2.868	0.720	2.806	0.730	5	5.745	0.332	6.122	0.295	5	3.373	0.643	3.379	0.642
6	3.299	0.771	3.304	0.770	6	7.570	0.271	7.878	0.247	6	3.947	0.684	4.062	0.668
7	3.699	0.814	3.716	0.812	7	7.877	0.344	8.010	0.332	7	4.296	0.745	4.453	0.726
8	5.421	0.712	5.312	0.724	8	8.774	0.362	8.542	0.382	8	6.135	0.632	6.304	0.613
9	5.468	0.792	5.334	0.804	9	8.822	0.454	8.596	0.475	9	6.390	0.700	6.447	0.695
10	7.097	0.716	6.660	0.757	10	10.100	0.432	9.855	0.453	10	8.822	0.549	8.523	0.578
Durbin-Watson				2.028	Durbin-Watson				2.050	Durbin-Watson				2.014
Pr<DW				0.673	Pr<DW				0.784	Pr<DW				0.585
Pr>DW				0.327	Pr>DW				0.216	Pr>DW				0.415
ID-BIG5	X				ID-BIG5	log V				ID-BIG5	T			
Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM
1	0.120	0.729	0.109	0.741	1	0.086	0.769	0.078	0.780	1	0.008	0.929	0.006	0.940
2	0.809	0.667	0.792	0.673	2	3.952	0.139	4.183	0.124	2	0.560	0.756	0.546	0.761
3	0.812	0.847	0.796	0.850	3	6.104	0.107	6.483	0.090	3	0.736	0.865	0.720	0.868
4	1.930	0.749	1.921	0.750	4	6.788	0.148	7.333	0.119	4	2.537	0.638	2.539	0.638
5	2.848	0.723	2.787	0.733	5	6.833	0.233	7.551	0.183	5	3.360	0.645	3.366	0.644
6	3.281	0.773	3.288	0.772	6	8.671	0.193	9.234	0.161	6	3.927	0.687	4.043	0.671
7	3.682	0.816	3.699	0.814	7	8.912	0.259	9.311	0.231	7	4.270	0.748	4.427	0.730
8	5.389	0.715	5.279	0.727	8	10.549	0.229	10.572	0.227	8	6.106	0.635	6.271	0.617
9	5.435	0.795	5.301	0.807	9	10.856	0.286	11.016	0.275	9	6.365	0.703	6.418	0.698
10	7.065	0.719	6.628	0.760	10	11.658	0.309	11.861	0.295	10	8.775	0.554	8.473	0.583
Durbin-Watson				2.033	Durbin-Watson				2.057	Durbin-Watson				2.018
Pr<DW				0.696	Pr<DW				0.816	Pr<DW				0.611
Pr>DW				0.304	Pr>DW				0.184	Pr>DW				0.389
ID-BIG5	log Wxv				ID-BIG5	Zxv				ID-BIG5	log Cxv			
Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM
1	0.029	0.866	0.034	0.854	1	0.107	0.744	0.097	0.756	1	0.024	0.878	0.021	0.886
2	2.751	0.253	2.844	0.241	2	0.795	0.672	0.778	0.678	2	4.963	0.084	4.978	0.083
3	5.135	0.162	5.520	0.138	3	0.802	0.849	0.787	0.853	3	5.630	0.131	5.690	0.128
4	5.652	0.227	6.149	0.188	4	1.943	0.746	1.933	0.748	4	6.223	0.183	6.483	0.166
5	5.662	0.341	6.269	0.281	5	2.852	0.723	2.795	0.732	5	6.345	0.274	6.757	0.239
6	8.297	0.217	8.781	0.186	6	3.298	0.771	3.313	0.769	6	8.453	0.207	8.683	0.192
7	8.304	0.307	8.788	0.268	7	3.700	0.814	3.728	0.811	7	8.826	0.265	8.870	0.262
8	8.733	0.365	8.992	0.343	8	5.419	0.712	5.326	0.722	8	9.326	0.316	9.084	0.335
9	9.156	0.423	9.524	0.390	9	5.473	0.791	5.352	0.803	9	9.336	0.407	9.109	0.427
10	10.095	0.432	10.365	0.409	10	7.139	0.712	6.710	0.753	10	10.483	0.399	10.053	0.436
Durbin-Watson				2.042	Durbin-Watson				2.031	Durbin-Watson				2.041
Pr<DW				0.749	Pr<DW				0.690	Pr<DW				0.737
Pr>DW				0.251	Pr>DW				0.310	Pr>DW				0.263

Table 3.4: DW, Q and LM tests for intraday model fits to GenBM data.

3.3.2 GenBIG Data

We now turn to the results of fitting the models of Section 3.2 to the GenBIG data, which was generated assuming underlying BIG innovation processes. Again, since the parameter values used to generate the data are known, we use this section for testing our programs and illustrating the underlying theory. Since the data is generated using an underlying BIG process we expect the models assuming NIG-distributed innovations (or underlying BIG processes) to fit better than those assuming normally distributed innovations (or underlying BM processes). Furthermore, the more intraday data is used in the models, the better they should fit.

ML estimates

Table 3.5 gives the ML estimates of the parameters, with asymptotic standard errors in brackets, of all the model fits to the GenBIG data. All the parameter estimates are within two standard errors of the true values and therefore (at least reasonably) accurate. The estimates generally become more accurate (with smaller standard errors) as more data is used, as well as when we move away from models assuming normally distributed innovations (or underlying BM processes) to models assuming NIG-distributed innovations (or underlying BIG processes). This illustrates the information gain as we include more data in the models as well as the fact that models assuming NIG-distributed innovations lead to more accurate parameter estimates than models assuming normally distributed innovations.

The fact that the D-NOR, HLC-BM and ID-BM models lead to reasonably accurate parameter estimates is in line with the consistency property of QML.

Quality of fit

From Figures 3.18 and 3.19 we see that the residuals and estimates of both the D-NOR and D-NIG models are close to their model-implied distributions, although the D-NIG model seems to fit the GenBIG data somewhat better, as we would expect.

Figures 3.20 and 3.21 show how well the distributions of the residuals and estimators resulting from the HLC-BM and HLC-BIG model fits to the GenBIG data correspond to their model implied distributions. Comparing the tails of the bottom panels of Figures 3.20 and 3.21, we see that the HLC-BIG model fits the GenBIG data better than the HLC-BM model, as expected. The distributions of $\hat{W}_{X,n}$ and $\hat{W}_{XAB,n}$ also do not differ much from their model implied distributions, as is shown in Figure 3.22, so that the HLC-BIG model fits the GenBIG data reasonably well.

Although the parameter estimates of the ID-BM model seem reasonable, we see from Figures 3.23 and 3.24 that the distribution of the $\hat{V}_n$ residuals obtained when fitting the ID-BM model to the

GenBIG data differ considerably from its model implied distribution. Thus, the ID-BM model does not fit the GenBIG data well, and the choice for $\psi = \infty$ must be discarded.

The ID-BIG model, however, does fit the GenBIG data well. Figures 3.25 and 3.26 show that the residuals and estimates obtained from fitting the ID-BIG model to the 5 minute GenBIG data do not differ substantially from their model implied distributions. The same results hold for the ID-BIG fit to the 30 minute GenBIG data, although some of the accuracy is lost due to the fact that less of the data is used in the model fit. Figures 3.27 and 3.28 show the corresponding results.

Parameter	GenBIG	D-NOR	D-NIG
β	0.20	0.1552 (0.3282)	0.0877 (0.1003)
ψ	2.00	-	2.1072 (0.4654)
ν	0.00	0.0008 (0.0035)	0.0016 (0.0010)
ϕ	0.15	0.1309 (0.0355)	0.1252 (0.0331)
$\alpha_0 x 10^5$	1.00	1.3473 (0.3928)	1.2293 (0.3481)
α_1	0.10	0.1312 (0.0259)	0.1322 (0.0259)
β_1	0.80	0.7602 (0.0373)	0.7714 (0.0366)
Parameter	**GenBIG**	**HLC-BM**	**HLC-BIG**
β	0.20	0.1291 (0.0846)	0.0970 (0.0685)
ψ	2.00	-	1.9192 (0.0922)
ν	0.00	0.0012 (0.0007)	0.0015 (0.0006)
ϕ	0.15	0.1618 (0.0323)	0.1350 (0.0290)
$\alpha_0 x 10^5$	1.00	0.8815 (0.1243)	0.9774 (0.1378)
α_1	0.10	0.0860 (0.0097)	0.1012 (0.0110)
β_1	0.80	0.8006 (0.0195)	0.7944 (0.0199)
Parameter	**GenBIG**	**ID-BM30**	**ID-BIG30**
β	0.20	0.1728 (0.0772)	0.1865 (0.0665)
ψ	2.00	-	1.9036 (0.0770)
ν	0.00	0.0008 (0.0007)	0.0007 (0.0006)
ϕ	0.15	0.1270 (0.0342)	0.1292 (0.0286)
$\alpha_0 x 10^5$	1.00	0.8519 (0.1272)	0.9367 (0.1385)
α_1	0.10	0.0917 (0.0094)	0.0907 (0.0096)
β_1	0.80	0.8112 (0.0186)	0.8035 (0.0201)
Parameter	**GenBIG**	**ID-BM5**	**ID-BIG5**
β	0.20	0.2193 (0.0706)	0.1962 (0.0545)
ψ	2.00	-	1.9807 (0.0587)
ν	0.00	0.0003 (0.0006)	0.0005 (0.0004)
ϕ	0.15	0.1126 (0.0438)	0.1244 (0.0254)
$\alpha_0 x 10^5$	1.00	1.0644 (0.1386)	1.1204 (0.1239)
α_1	0.10	0.0991 (0.0090)	0.1003 (0.0085)
β_1	0.80	0.7911 (0.0190)	0.7853 (0.0170)

Table 3.5: ML estimates of all model fits to GenBIG data.

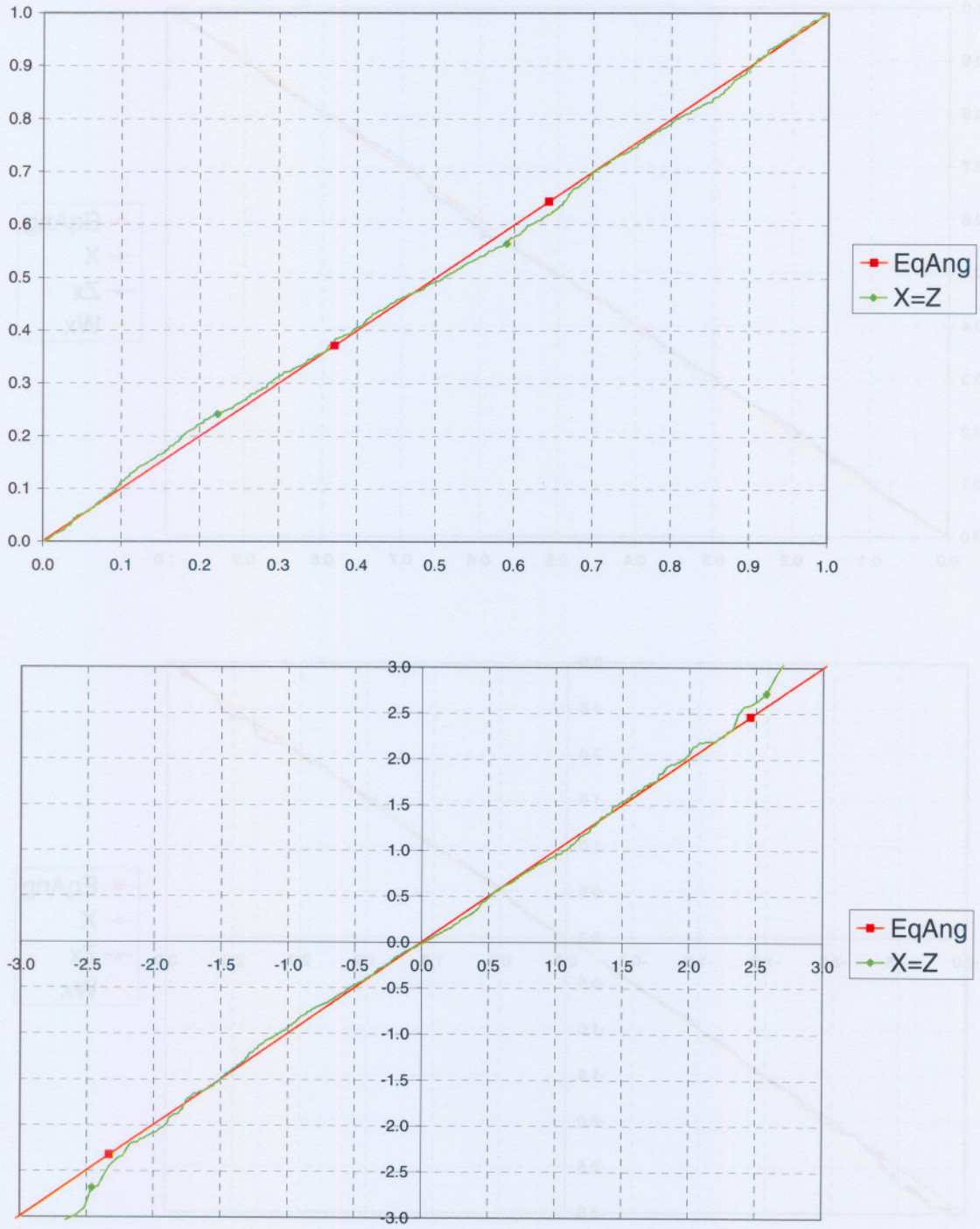

Figure 3.18: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$ and $\hat{Z}_n$ checking D-NOR fit to GenBIG data.

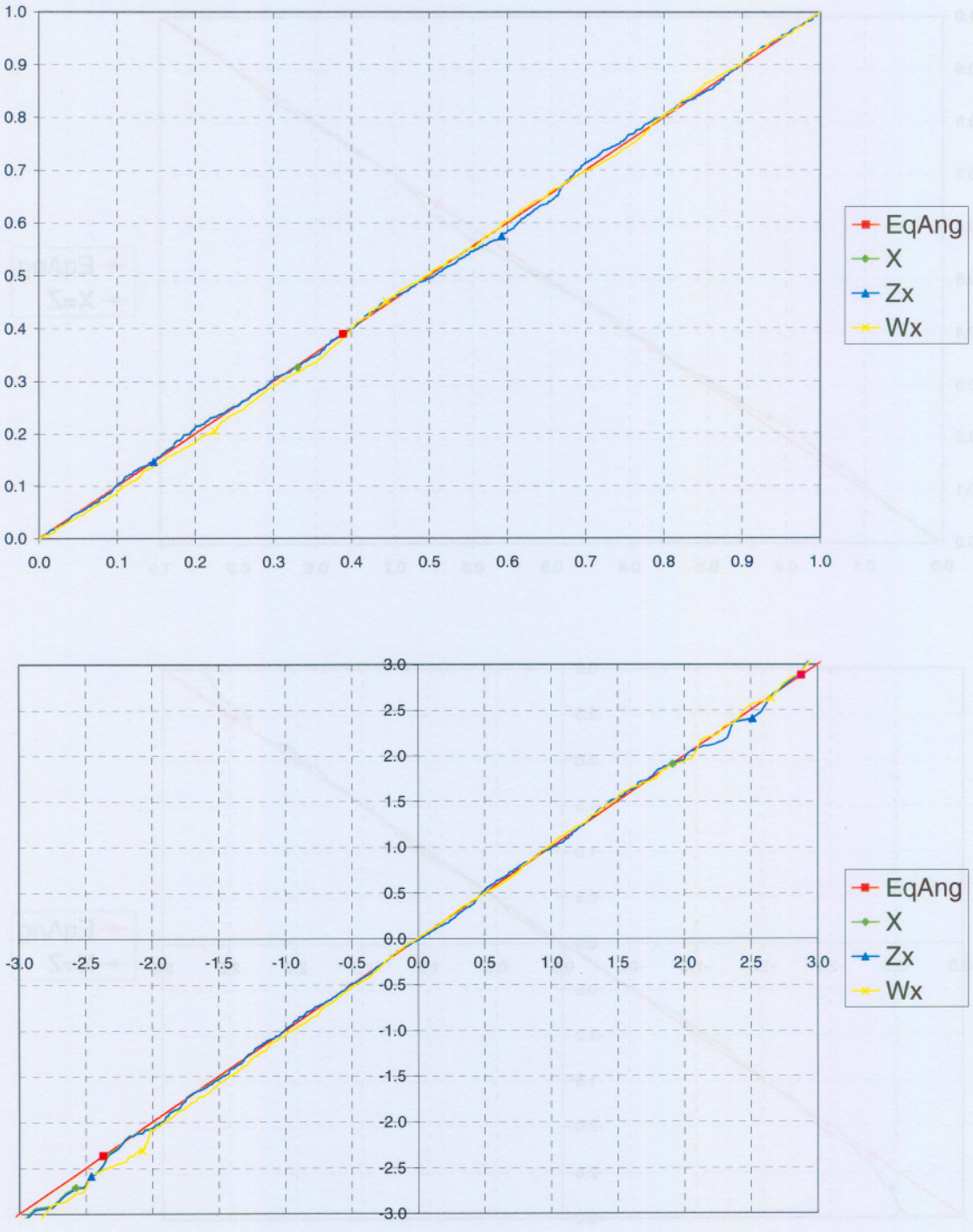

Figure 3.19: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_{X,n}$ and $\hat{W}_{X,n}$ checking D-NIG fit to GenBIG data.

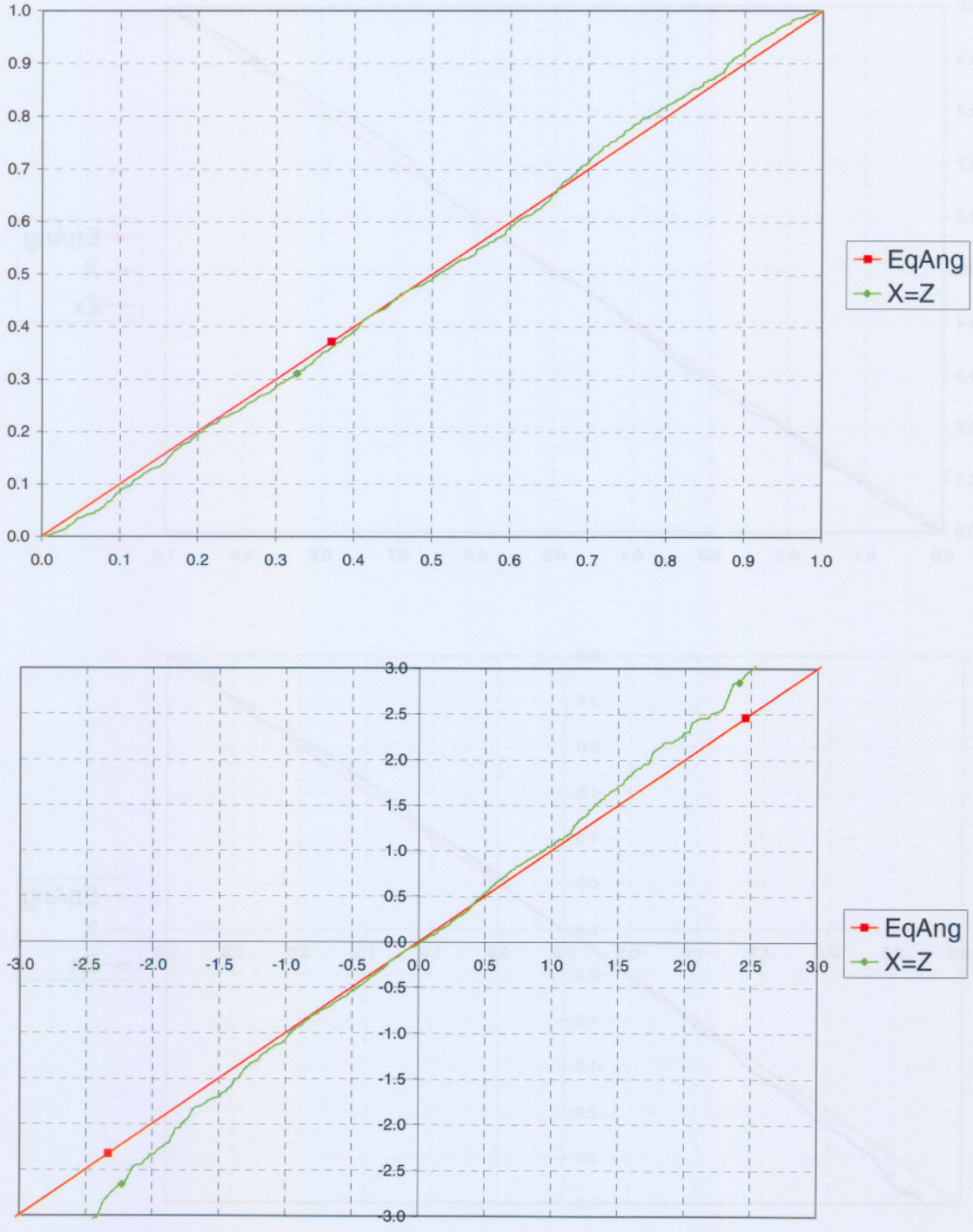

Figure 3.20: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$ and $\hat{Z}_n$ checking HLC-BM fit to GenBIG data.

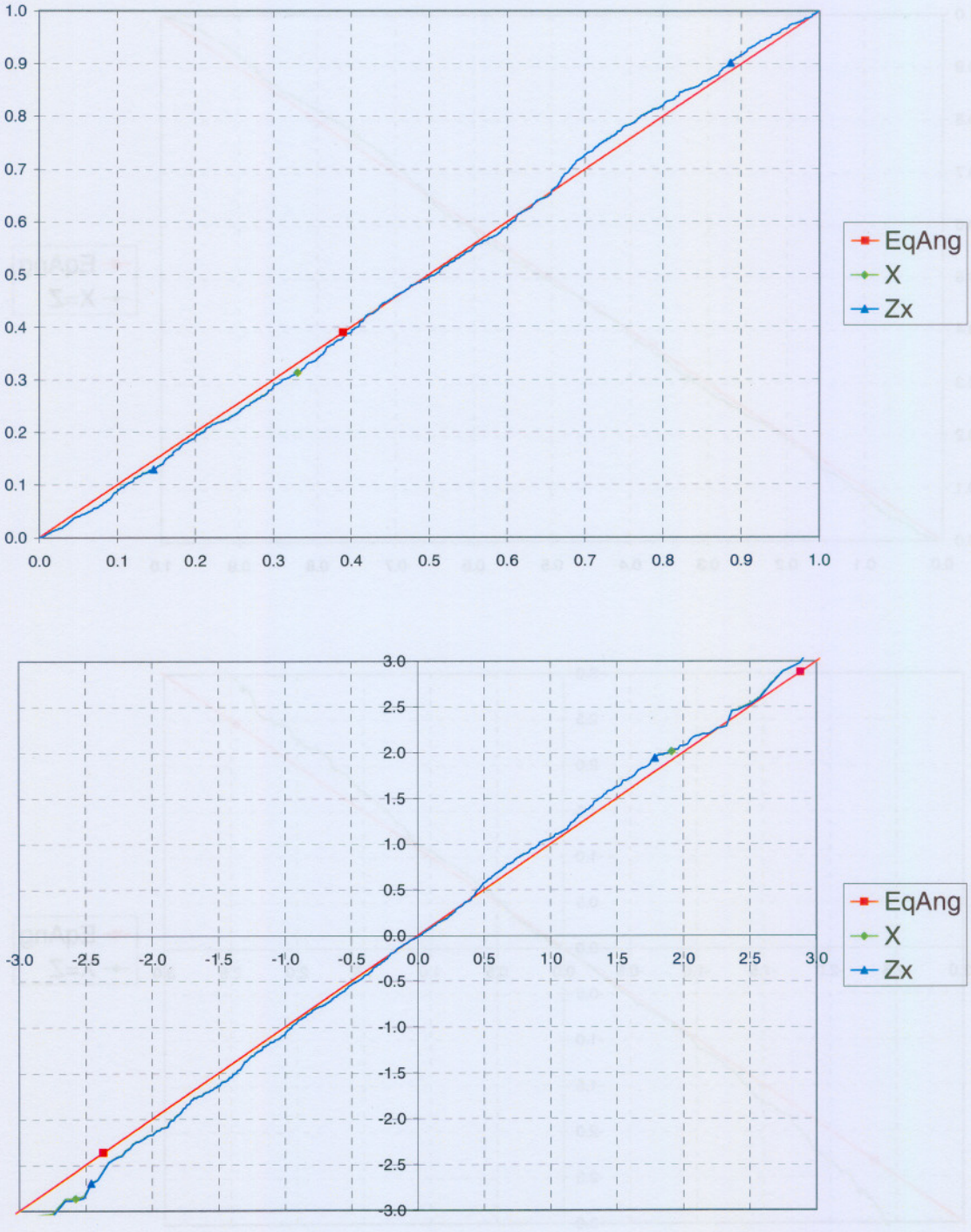

Figure 3.21: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$ and $\hat{Z}_{X,n}$ checking HLC-BIG fit to GenBIG data.

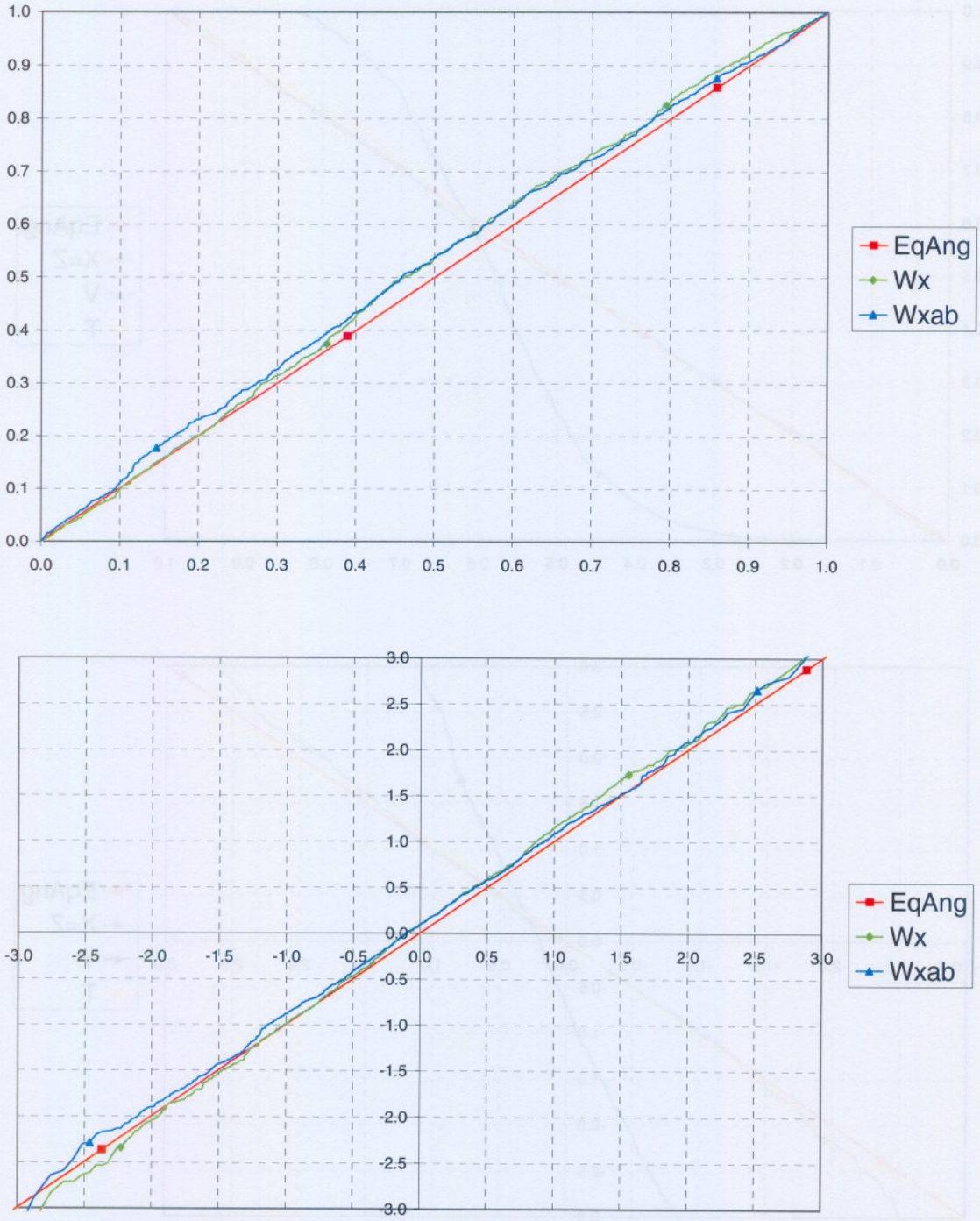

Figure 3.22: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{X,n}$ and $\hat{W}_{XAB,n}$ checking HLC-BIG fit to GenBIG data.

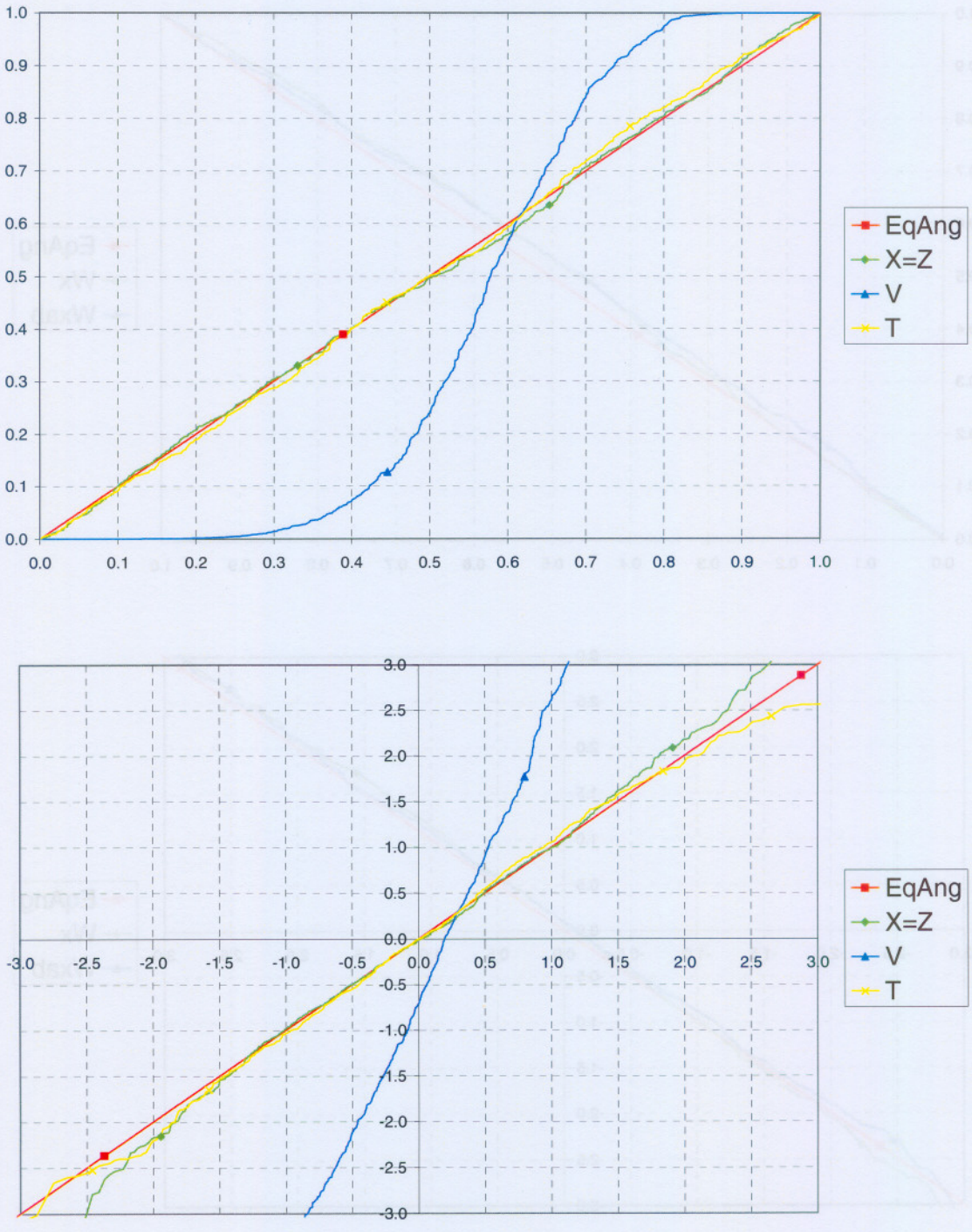

Figure 3.23: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BM5 fit to GenBIG data.

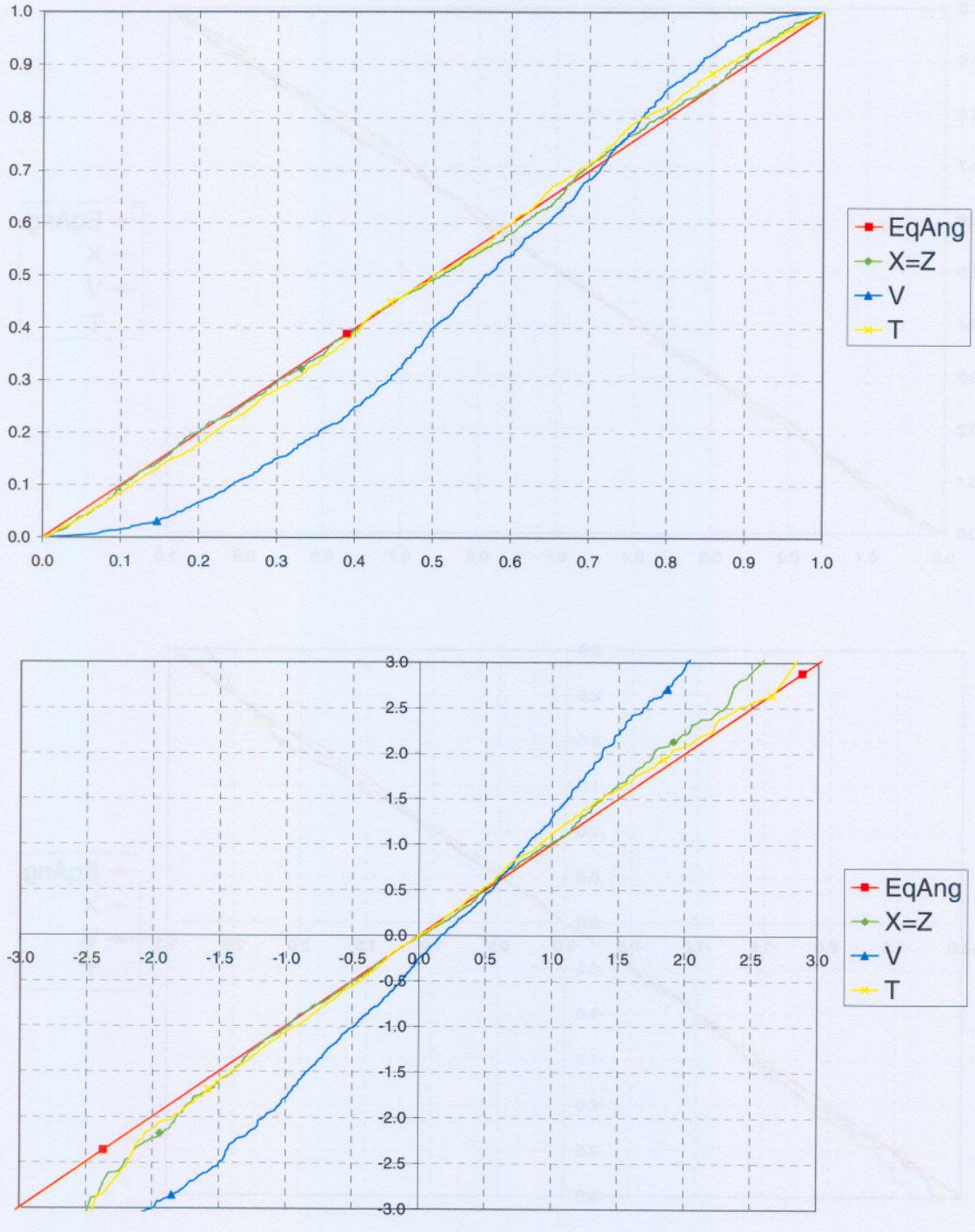

Figure 3.24: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BM30 fit to GenBIG data.

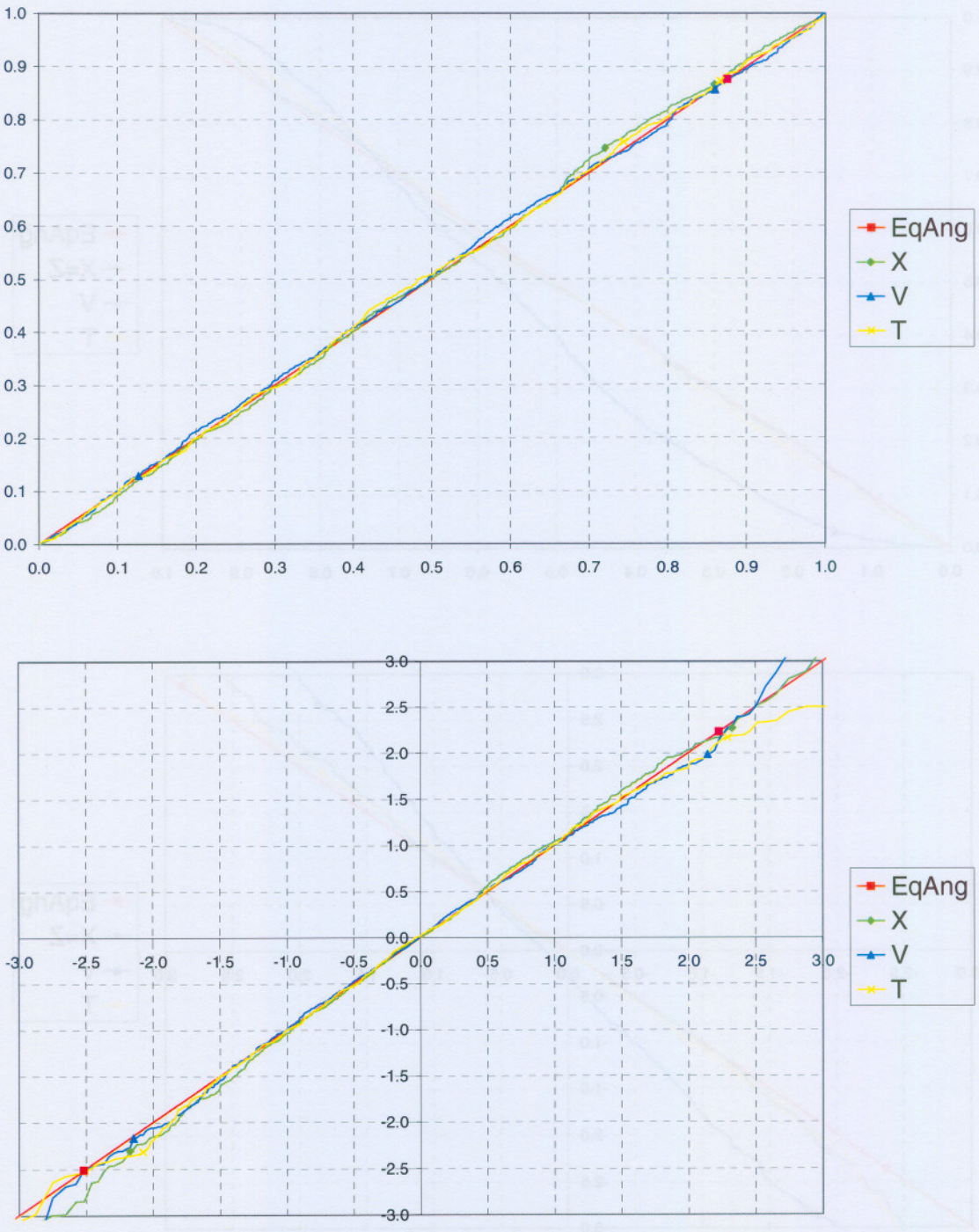

Figure 3.25: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BIG5 fit to GenBIG data.

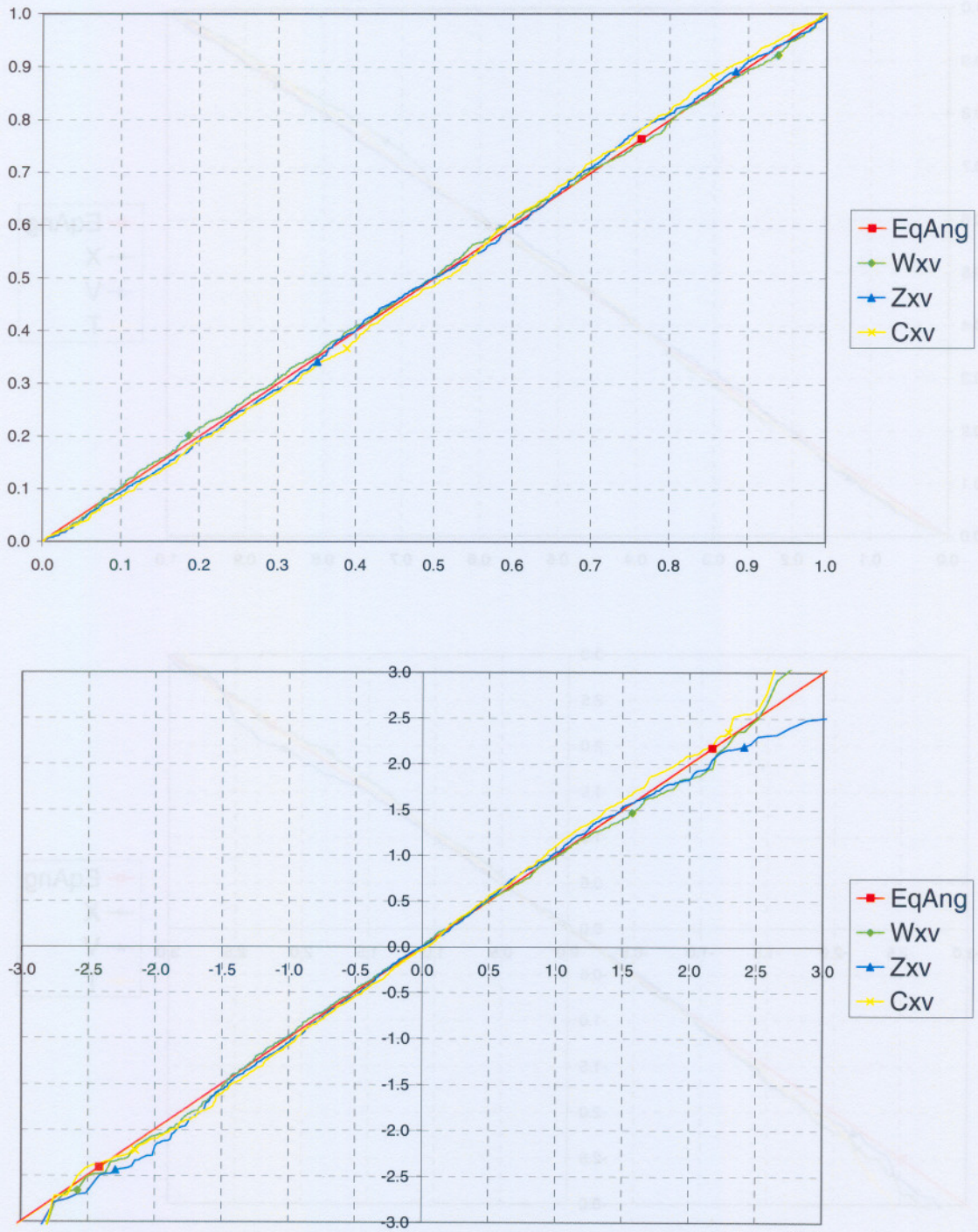

Figure 3.26: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{XV,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking ID-BIG5 fit to GenBIG data.

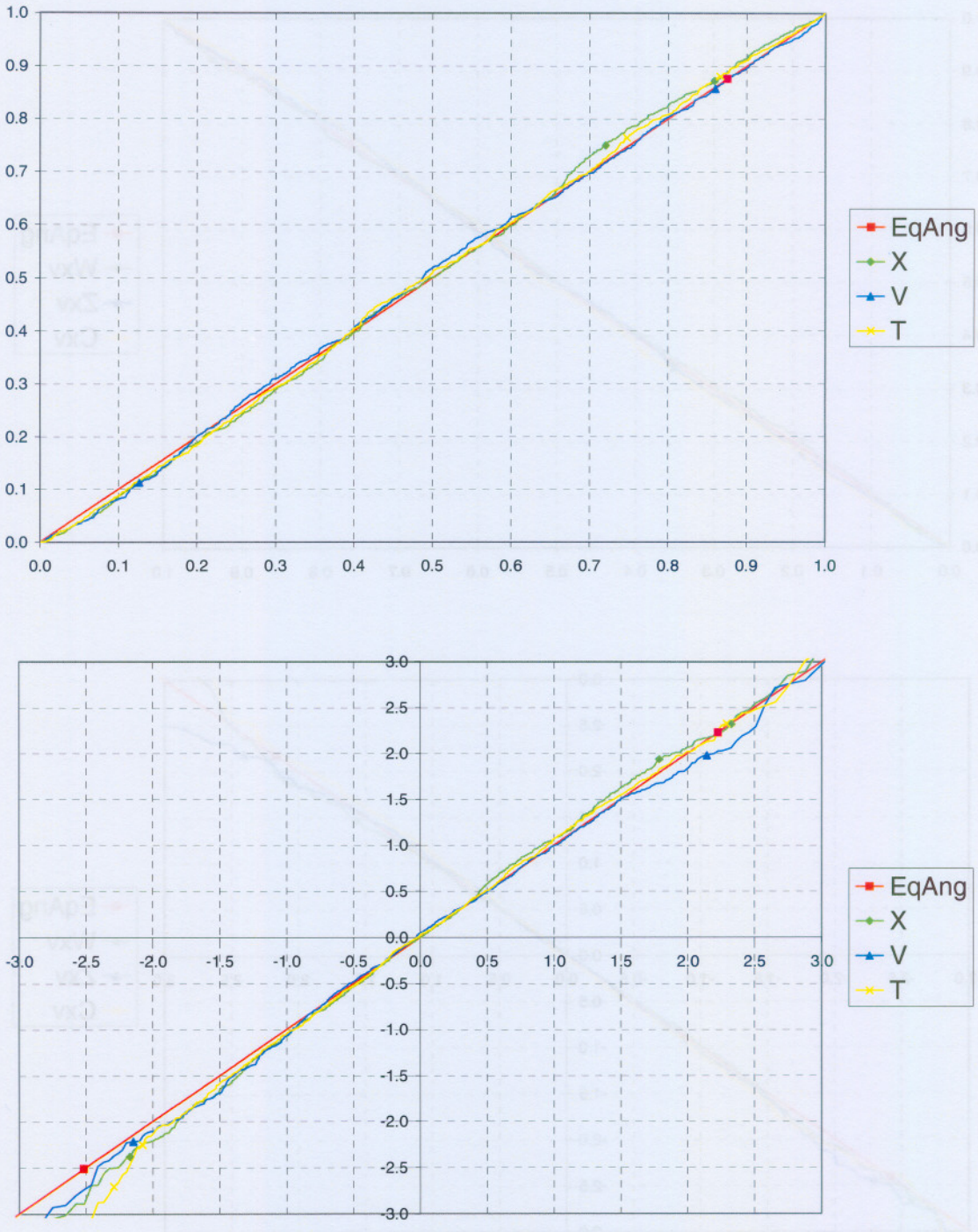

Figure 3.27: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BIG30 fit to GenBIG data.

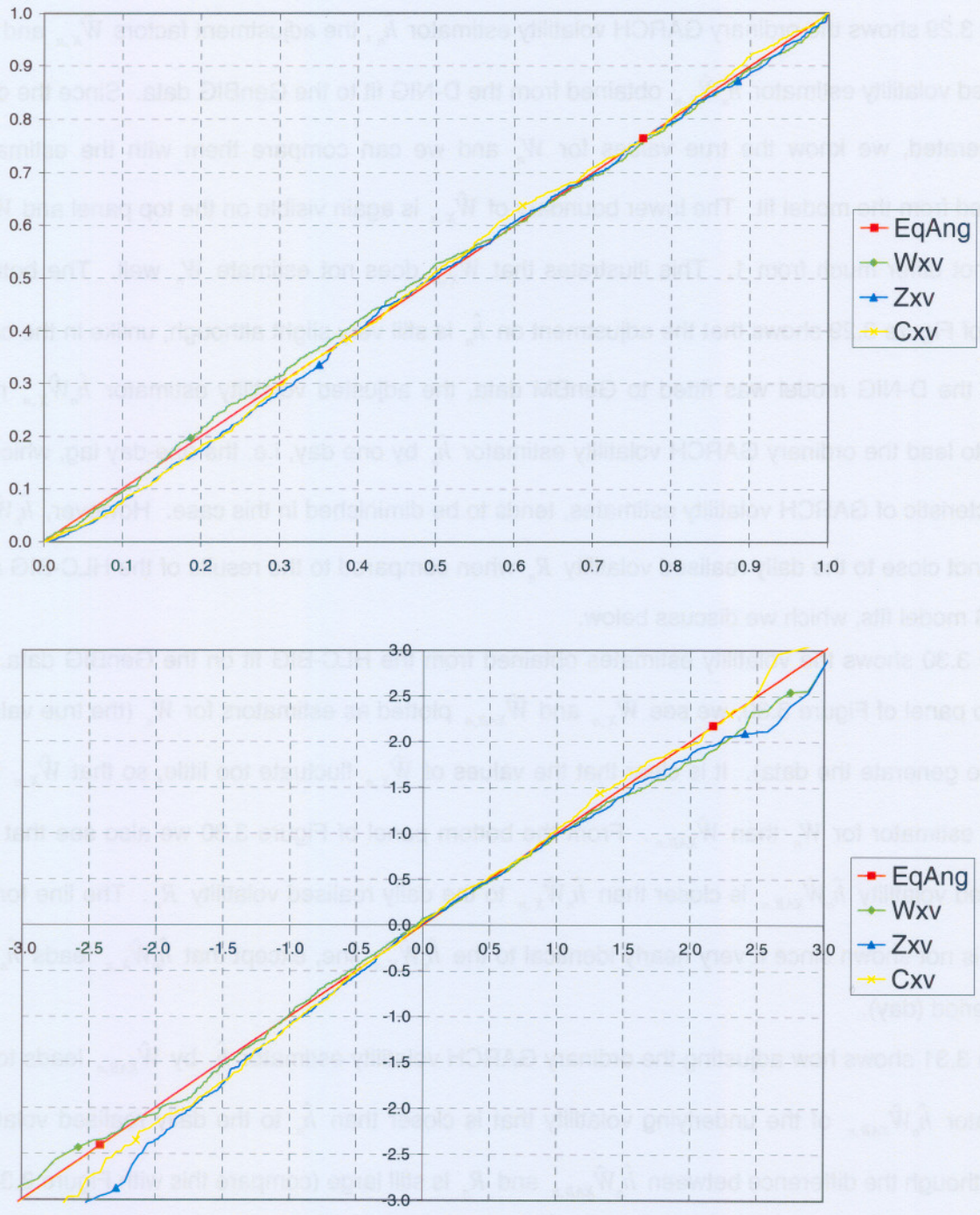

Figure 3.28: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{XV,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking ID-BIG30 fit to GenBIG data.

Volatility estimates

Figure 3.29 shows the ordinary GARCH volatility estimator $\hat{h}_n$, the adjustment factors $\hat{W}_{X,n}$ and the adjusted volatility estimator $\hat{h}_n\hat{W}_{X,n}$ obtained from the D-NIG fit to the GenBIG data. Since the data is generated, we know the true values for W_n and we can compare them with the estimates obtained from the model fit. The lower boundary of $\hat{W}_{X,n}$ is again visible on the top panel and $\hat{W}_{X,n}$ does not differ much from 1. This illustrates that $\hat{W}_{X,n}$ does not estimate W_n well. The bottom panel of Figure 3.29 shows that the adjustment on $\hat{h}_n$ is still very slight although, unlike in the case where the D-NIG model was fitted to GenBM data, the adjusted volatility estimator $\hat{h}_n\hat{W}_{X,n}$ now tends to lead the ordinary GARCH volatility estimator $\hat{h}_n$ by one day, i.e. the one-day lag, which is characteristic of GARCH volatility estimates, tends to be diminished in this case. However, $\hat{h}_n\hat{W}_{X,n}$ is still not close to the daily realised volatility R_n when compared to the results of the HLC-BIG and ID-BIG model fits, which we discuss below.

Figure 3.30 shows the volatility estimates obtained from the HLC-BIG fit on the GenBIG data. In the top panel of Figure 3.30, we see $\hat{W}_{X,n}$ and $\hat{W}_{XAB,n}$ plotted as estimators for W_n (the true values used to generate the data). It is clear that the values of $\hat{W}_{X,n}$ fluctuate too little, so that $\hat{W}_{X,n}$ is a worse estimator for W_n than $\hat{W}_{XAB,n}$. From the bottom panel of Figure 3.30 we also see that the adjusted volatility $\hat{h}_n\hat{W}_{XAB,n}$ is closer than $\hat{h}_n\hat{W}_{X,n}$ to the daily realised volatility R_n. The line for $\hat{h}_n$ alone is not shown since it very nearly identical to the $\hat{h}_n\hat{W}_{X,n}$ line, except that $\hat{h}_n\hat{W}_{X,n}$ leads $\hat{h}_n$ by one period (day).

Figure 3.31 shows how adjusting the ordinary GARCH volatility estimator $\hat{h}_n$ by $\hat{W}_{XAB,n}$ leads to an estimator $\hat{h}_n\hat{W}_{XAB,n}$ of the underlying volatility that is closer than $\hat{h}_n$ to the daily realised volatility R_n, although the difference between $\hat{h}_n\hat{W}_{XAB,n}$ and R_n is still large (compare this with Figure 3.33).

Figure 3.29: Plots of W_n and $\hat{W}_{X,n}$ (top panel) and R_n, $\hat{h}_n$, and $\hat{h}_n\hat{W}_{X,n}$ (bottom panel) for D-NIG fit to GenBIG data.

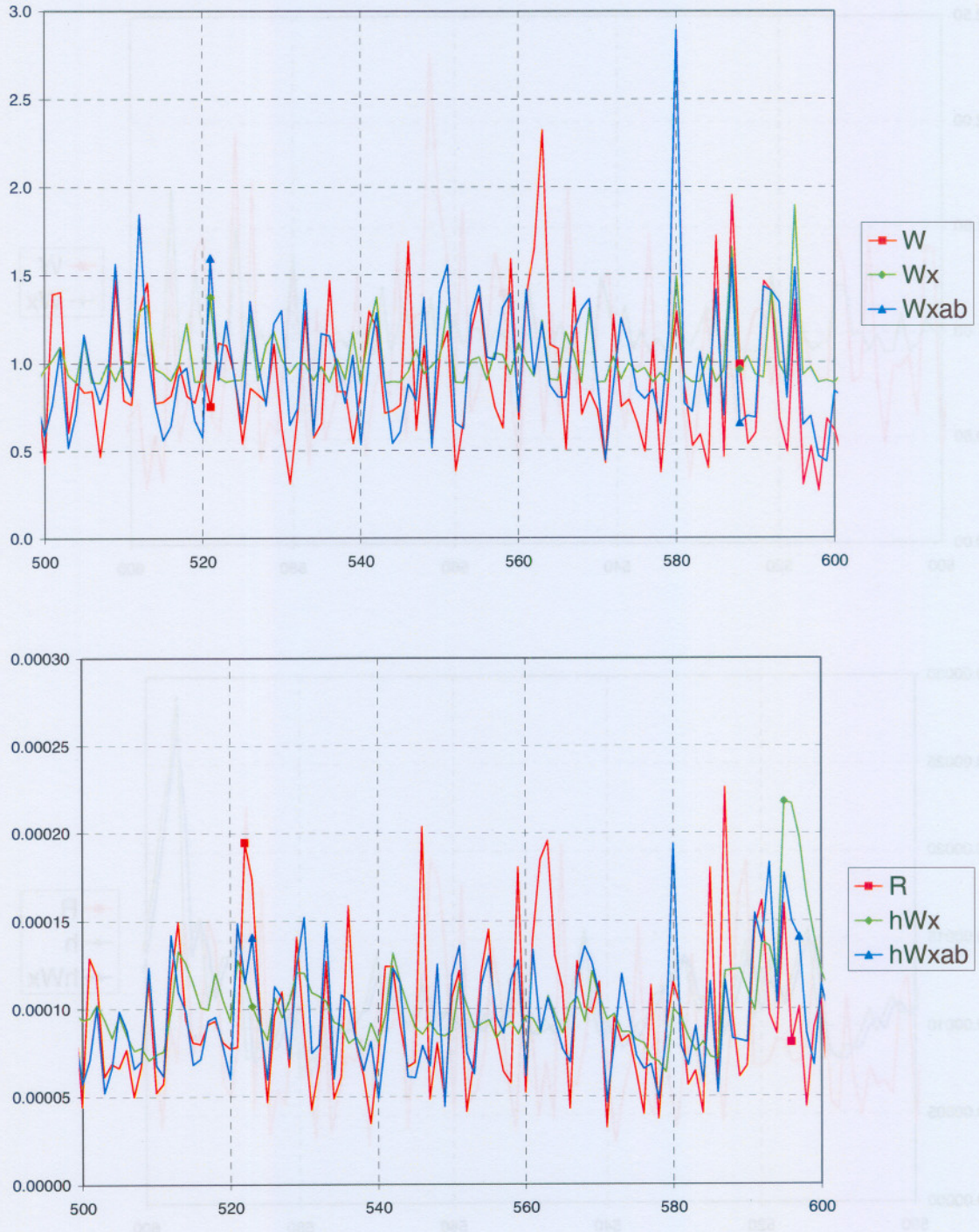

Figure 3.30: Plots of W_n, $\hat{W}_{X,n}$ and $\hat{W}_{XAB,n}$ (top panel) and R_n, $\hat{h}_n\hat{W}_{X,n}$ and $\hat{h}_n\hat{W}_{XAB,n}$ (bottom panel) for HLC-BIG fit to GenBIG data.

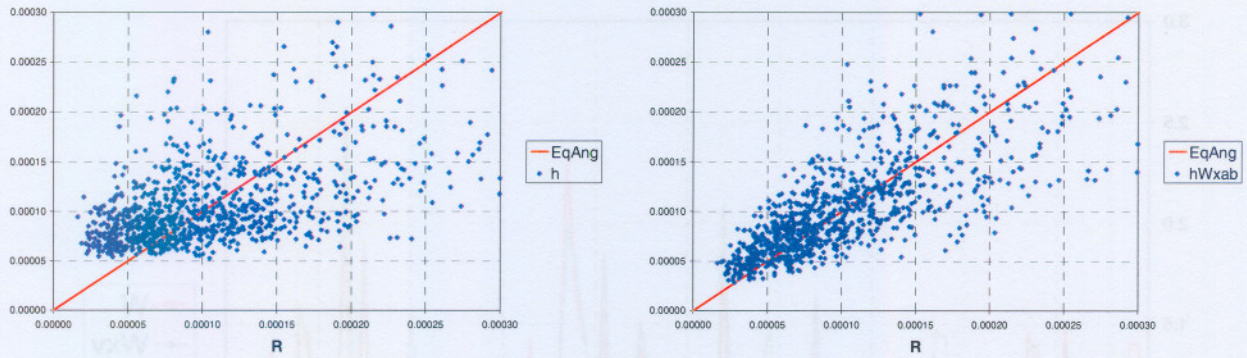

Figure 3.31: Plots of $\hat{h}_n$ (left) and $\hat{h}_n\hat{W}_{XAB,n}$ (right) against R_n for HLC-BIG fit to GenBIG data.

Moving on to the volatility estimates obtained from the ID-BIG model fit to the GenBIG data, the top panel of Figure 3.32 shows that $\hat{W}_{XV,n}$ estimates W_n very accurately. Again, note that when using real-world data, the true values of the latent variables W_n will not be known, so that these kinds of comparisons will not be possible. We show these comparisons in order to test the model fit when the true values are indeed known. The bottom panel of Figure 3.32 shows that the adjusted estimator $\hat{h}_n\hat{W}_{XV,n}$ is much closer to the daily realised volatility R_n than the ordinary GARCH volatility estimator $\hat{h}_n$. This can also be seen in Figure 3.33, where $\hat{h}_n$ and $\hat{h}_n\hat{W}_{XV,n}$ are plotted against R_n. Note that the slope of $\hat{h}_n\hat{W}_{XV,n}$ against R_n is slightly less than 1. When considering (3.2.18), and the argument following it, this can be explained. We have already argued that, since $C_n \sim \chi^2_{I-1}$, $C_n/I \to 1$ as $I \to \infty$ so that $S_n \approx \hat{h}_n\hat{W}_n$, and since (2.7.14) states that $S_n = R_n - Y_1^2/I$, it follows that $R_n \approx S_n \approx \hat{h}_n\hat{W}_n$. However, $E(C_n/I) = (I-1)/I$, so that the slope of $\hat{h}_n\hat{W}_{XV,n}$ against R_n should become increasingly smaller than 1 for smaller values of I. The influence of Y_1^2/I also becomes greater for smaller values of I. These two factors cause R_n to be increasingly larger than $\hat{h}_n\hat{W}_n$ for smaller values of I. In this case $I = 78$ is quite large so that we would expect the slope to be only slightly less than 1 and this is indeed shown in Figure 3.33.

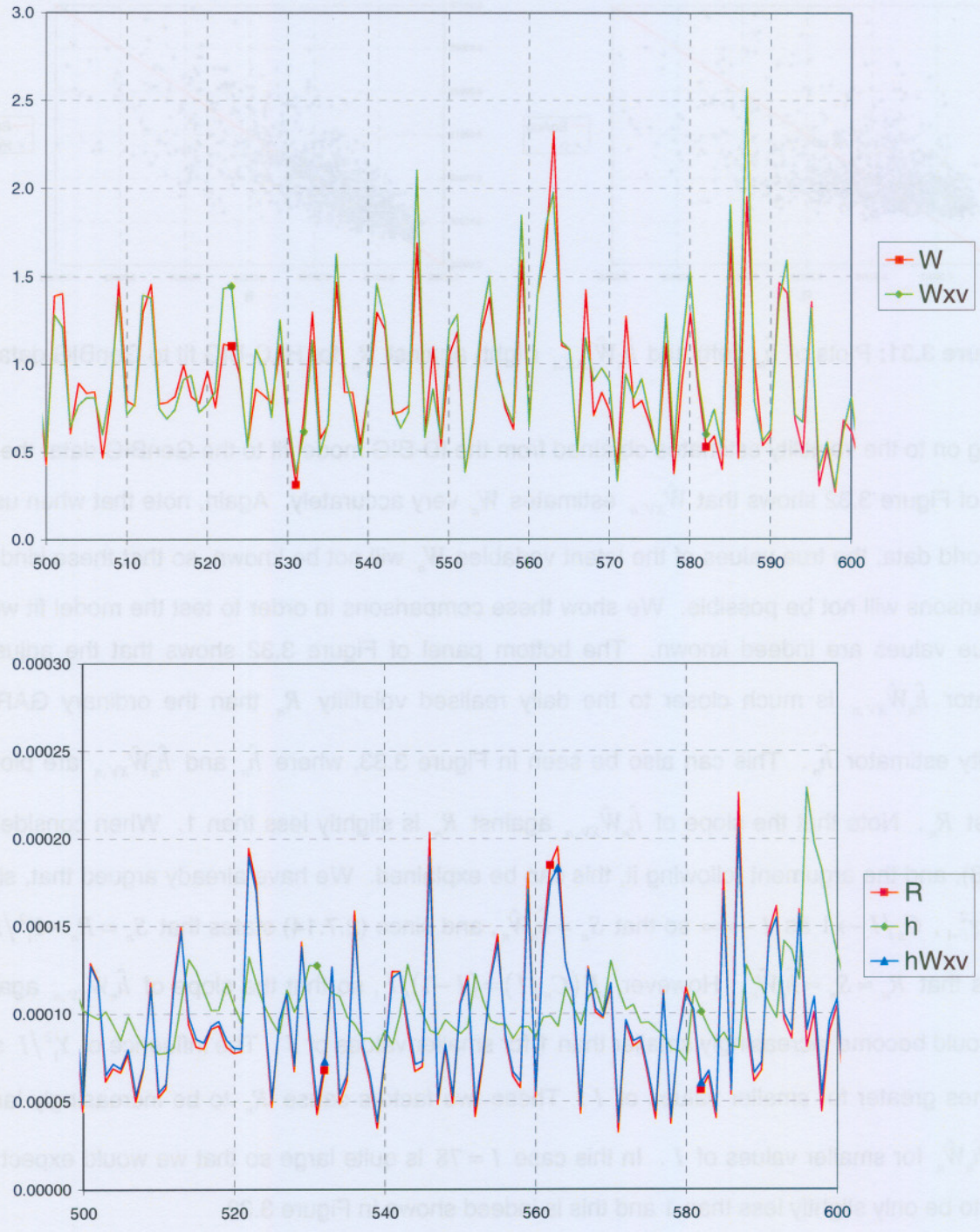

Figure 3.32: Plots of W_n and $\hat{W}_{XV,n}$ (top panel) and R_n, $\hat{h}_n$ and $\hat{h}_n\hat{W}_{XV,n}$ (bottom panel) for ID-BIG5 fit to GenBIG data.

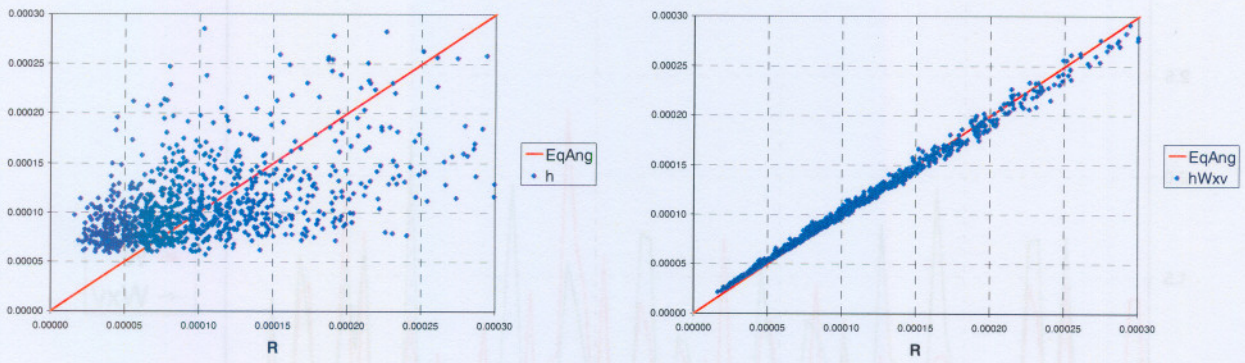

Figure 3.33: Plots of $\hat{h}_n$ (left) and $\hat{h}_n\hat{W}_{XV,n}$ (right) against R_n for ID-BIG5 fit to GenBIG data.

Comparing the volatility estimates obtained from the ID-BIG model fit to the 5 minute data with those obtained from the ID-BIG model fit to the 30 minute data we see from the top panel of Figure 3.34 that $\hat{W}_{XV,n}$ still estimates W_n fairly accurately in the 30 minute case, although not as accurately as in the 5 minute case, as expected. The adjusted volatility estimator $\hat{h}_n\hat{W}_{XV,n}$ is also much closer than the ordinary GARCH volatility estimator $\hat{h}_n$ to the realised volatility R_n, as can be seen from the bottom panel of Figure 3.34, as well as Figure 3.35. In this case, since $I=13$, we would expect the slope of $\hat{h}_n\hat{W}_{XV,n}$ to be somewhat less than in the 5 minute case, and indeed the right-hand panel of Figure 3.35 confirms this.

Testing for dependence

Tables 3.6 and 3.7 contain the results of the DW, Q and LM tests for the model fits to the GenBIG data (as argued before we perform the tests on the natural logarithm of $\hat{W}_{X,n}$, $\hat{W}_{XAB,n}$, $\hat{W}_{XV,n}$, $\hat{V}_n$ and $\hat{C}_{XV,n}$). The Q and LM tests give no indication of remaining ARCH effects for any of the model fits, while the DW tests give no indication of autocorrelation in any of the residuals and estimates except for $\hat{X}_n$, $\hat{Z}_n$ and $\hat{T}_n$ obtained from the ID-BM model fit. We have already seen above that the ID-BM model does not fit the GenBIG data well. Tables 3.6 and 3.7 indicate no irregularities for any of the other models.

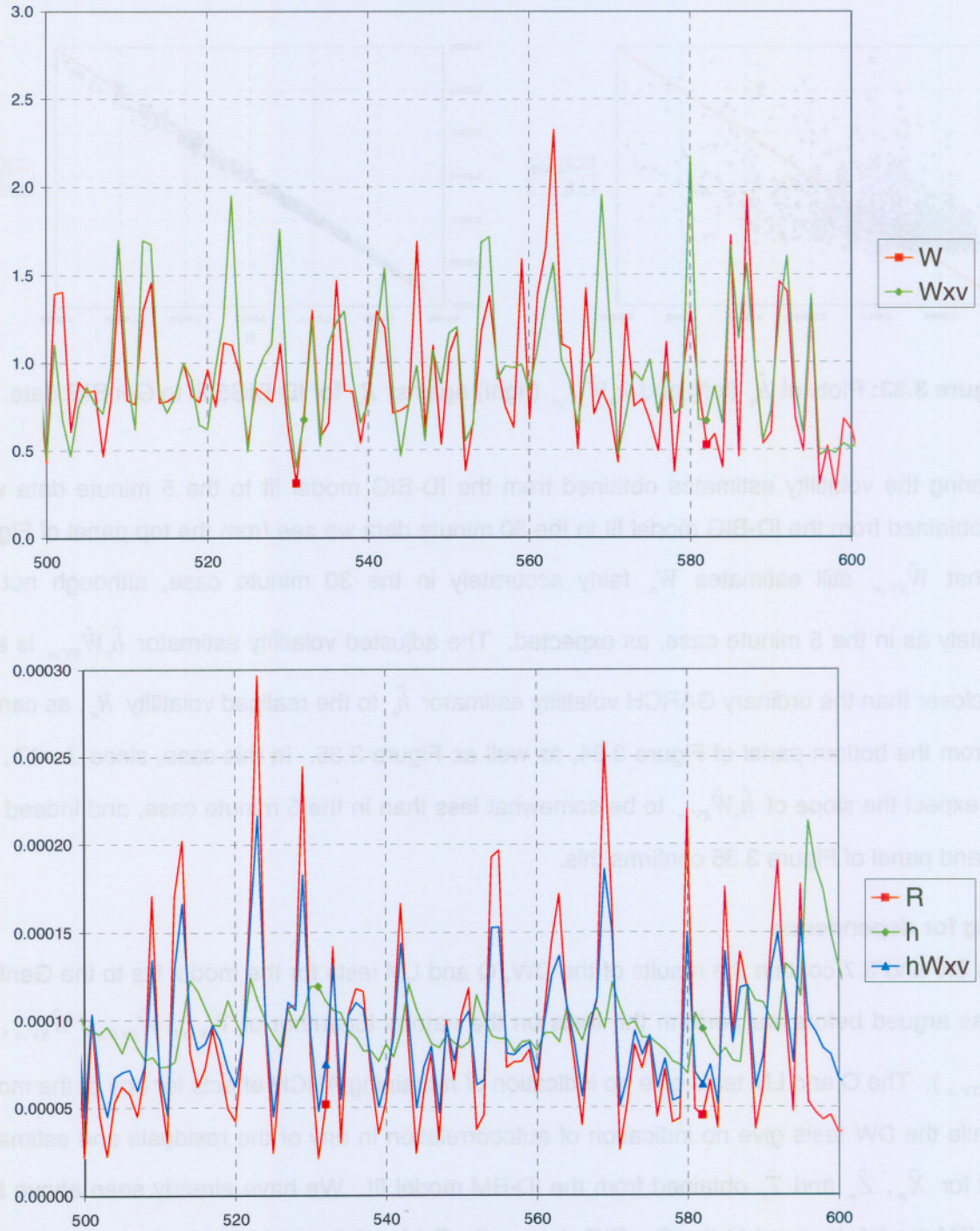

Figure 3.34: Plots of W_n and $\hat{W}_{XV,n}$ (top panel) and R_n, $\hat{h}_n$ and $\hat{h}_n\hat{W}_{XV,n}$ (bottom panel) for ID-BIG30 fit to GenBIG data.

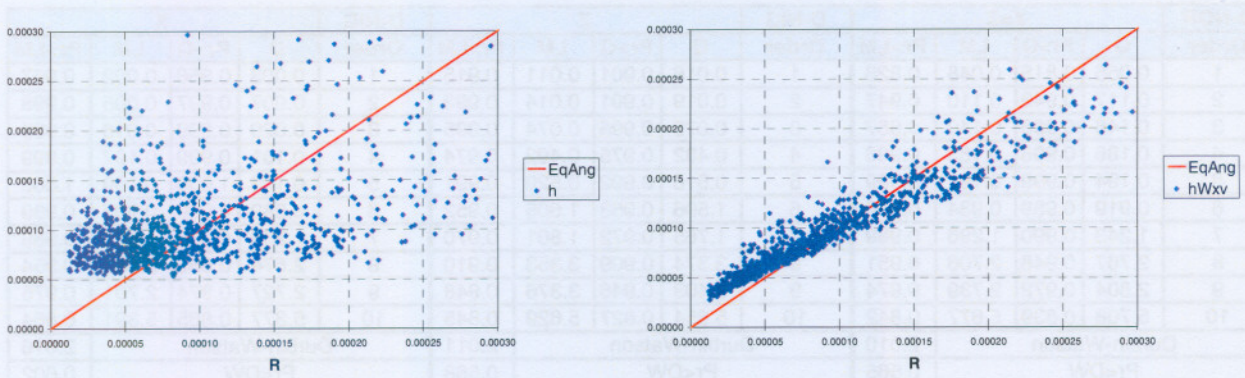

Figure 3.35: Plots of $\hat{h}_n$ (left) and $\hat{h}_n \hat{W}_{XV,n}$ (right) against R_n for ID-BIG30 fit to GenBIG data.

The case $\beta = 0$

Finally, consider what happens when we choose $\beta = 0$ in the models discussed above. Figure 3.36 shows the PIT plots and Q-Q plots for the $\hat{X}_n$ residuals when fitting the ID-BIG5 model to the GenBIG data. The plots shown for $\beta \neq 0$ are the same plots for $\hat{X}_n$ as shown in Figure 3.25. When we choose $\beta = 0$, the values and standard errors for the other parameter estimates change slightly, but the estimates are still fairly accurate and therefore not shown here. However, we can see in Figure 3.36 that the distribution of the $\hat{X}_n$ residuals when $\beta = 0$ differ slightly more from its model implied distribution than when $\beta \neq 0$, making the full model with $\beta \neq 0$ more appropriate. Furthermore, this effect is more prominent when fitting the ID-BM model (and other models) to the GenBIG data with $\beta = 0$.

We have seen that the models based on NIG-distributed innovations (or underlying BIG processes) fit the GenBIG data better than models based on normally distributed innovations (or underlying BM processes). Also, the parameter estimates and the estimates for the latent variables become more accurate as more data is used in the model fits. The ID-BIG5 model provides the best fit to the GenBIG data, as we would expect since the GenBIG data was generated by the processes underlying the ID-BIG model and the ID-BIG5 model uses as much of the available data as possible.

D-NOR	X=Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.055	0.815	0.048	0.826
2	0.117	0.943	0.110	0.947
3	0.146	0.986	0.140	0.987
4	0.186	0.996	0.183	0.996
5	0.194	0.999	0.191	0.999
6	0.919	0.989	0.934	0.988
7	1.243	0.990	1.296	0.989
8	2.767	0.948	2.706	0.951
9	2.804	0.972	2.739	0.974
10	5.708	0.839	5.677	0.842
Durbin-Watson				2.010
Pr<DW				0.565
Pr>DW				0.435

D-NIG	Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.016	0.901	0.011	0.915
2	0.019	0.991	0.014	0.993
3	0.075	0.995	0.074	0.995
4	0.482	0.975	0.492	0.974
5	0.516	0.992	0.529	0.991
6	1.596	0.953	1.605	0.952
7	1.765	0.972	1.801	0.970
8	3.374	0.909	3.353	0.910
9	3.403	0.946	3.376	0.948
10	5.854	0.827	5.629	0.845
Durbin-Watson				2.011
Pr<DW				0.568
Pr>DW				0.432

D-NIG	X			
Order	Q	Pr>Q	LM	Pr>LM
1	0.003	0.956	0.002	0.967
2	0.007	0.997	0.005	0.998
3	0.020	0.999	0.018	0.999
4	0.101	0.999	0.107	0.999
5	0.103	1.000	0.109	1.000
6	0.889	0.990	0.907	0.989
7	1.236	0.990	1.279	0.989
8	2.674	0.953	2.653	0.954
9	2.727	0.974	2.701	0.975
10	5.377	0.865	5.391	0.864
Durbin-Watson				2.016
Pr<DW				0.602
Pr>DW				0.398

D-NIG	log Wx			
Order	Q	Pr>Q	LM	Pr>LM
1	0.732	0.392	0.725	0.395
2	0.819	0.664	0.799	0.671
3	0.973	0.808	0.940	0.816
4	3.067	0.547	2.836	0.586
5	3.070	0.689	2.836	0.725
6	3.075	0.799	2.837	0.829
7	3.330	0.853	3.113	0.874
8	3.699	0.883	3.328	0.912
9	4.262	0.893	3.902	0.918
10	4.818	0.903	4.622	0.915
Durbin-Watson				2.054
Pr<DW				0.802
Pr>DW				0.198

HLC-BM	X=Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.299	0.585	0.284	0.594
2	0.818	0.664	0.808	0.668
3	0.911	0.823	0.915	0.822
4	0.919	0.922	0.924	0.921
5	0.954	0.966	0.966	0.965
6	1.364	0.968	1.372	0.968
7	1.534	0.981	1.571	0.980
8	3.096	0.928	2.959	0.937
9	3.181	0.957	3.031	0.963
10	5.994	0.816	5.830	0.829
Durbin-Watson				2.095
Pr<DW				0.067
Pr>DW				0.933

HLC-BIG	Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.495	0.482	0.465	0.495
2	0.961	0.618	0.947	0.623
3	1.145	0.766	1.163	0.762
4	1.291	0.863	1.296	0.862
5	1.483	0.915	1.497	0.914
6	2.223	0.898	2.228	0.898
7	2.297	0.942	2.334	0.939
8	4.028	0.855	3.866	0.869
9	4.049	0.908	3.879	0.919
10	6.746	0.749	6.280	0.791
Durbin-Watson				2.049
Pr<DW				0.781
Pr>DW				0.219

HLC-BIG	X			
Order	Q	Pr>Q	LM	Pr>LM
1	0.332	0.565	0.316	0.574
2	0.801	0.670	0.790	0.674
3	0.899	0.826	0.903	0.825
4	0.900	0.925	0.905	0.924
5	0.948	0.967	0.960	0.966
6	1.410	0.965	1.420	0.965
7	1.608	0.978	1.654	0.977
8	3.232	0.919	3.094	0.928
9	3.289	0.952	3.141	0.958
10	6.255	0.793	6.096	0.807
Durbin-Watson				2.057
Pr<DW				0.817
Pr>DW				0.183

HLC-BIG	log Wx			
Order	Q	Pr>Q	LM	Pr>LM
1	0.494	0.482	0.487	0.485
2	0.961	0.619	0.931	0.628
3	1.040	0.792	0.994	0.803
4	4.248	0.374	3.919	0.417
5	4.248	0.514	3.920	0.561
6	4.307	0.635	3.929	0.686
7	4.542	0.716	4.199	0.757
8	4.911	0.767	4.376	0.822
9	5.530	0.786	5.022	0.832
10	6.100	0.807	5.804	0.831
Durbin-Watson				2.031
Pr<DW				0.686
Pr>DW				0.314

HLC-BIG	log Wxab			
Order	Q	Pr>Q	LM	Pr>LM
1	0.314	0.575	0.308	0.579
2	1.651	0.438	1.704	0.427
3	2.754	0.431	2.837	0.418
4	3.477	0.481	3.307	0.508
5	8.251	0.143	8.134	0.149
6	9.139	0.166	9.068	0.170
7	9.170	0.241	9.095	0.246
8	11.770	0.162	11.417	0.179
9	11.921	0.218	11.566	0.239
10	12.779	0.236	11.798	0.299
Durbin-Watson				2.026
Pr<DW				0.657
Pr>DW				0.343

Table 3.6: DW, Q and LM tests for daily and HLC model fits to GenBIG data.

ID-BM5	X=Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.039	0.843	0.034	0.853
2	0.578	0.749	0.564	0.754
3	0.586	0.900	0.572	0.903
4	0.678	0.954	0.662	0.956
5	0.722	0.982	0.705	0.983
6	1.010	0.985	0.997	0.986
7	1.062	0.994	1.069	0.994
8	2.438	0.965	2.301	0.970
9	2.477	0.982	2.340	0.985
10	4.724	0.909	4.569	0.918
Durbin-Watson				2.225
Pr<DW				0.999
Pr>DW				0.001

ID-BM5	log V			
Order	Q	Pr>Q	LM	Pr>LM
1	0.116	0.734	0.121	0.728
2	0.250	0.883	0.261	0.878
3	1.612	0.657	1.623	0.654
4	3.216	0.522	3.074	0.546
5	3.245	0.662	3.098	0.685
6	3.262	0.775	3.121	0.794
7	5.840	0.559	5.450	0.605
8	5.945	0.653	5.657	0.686
9	7.926	0.542	7.535	0.582
10	8.814	0.550	8.069	0.622
Durbin-Watson				2.025
Pr<DW				0.656
Pr>DW				0.344

ID-BM5	T			
Order	Q	Pr>Q	LM	Pr>LM
1	0.189	0.664	0.201	0.654
2	1.707	0.426	1.715	0.424
3	1.812	0.612	1.794	0.616
4	2.490	0.647	2.464	0.651
5	2.996	0.701	3.016	0.698
6	3.298	0.771	3.436	0.753
7	3.497	0.836	3.622	0.822
8	5.126	0.744	5.146	0.742
9	5.489	0.790	5.372	0.801
10	6.761	0.748	6.420	0.779
Durbin-Watson				2.212
Pr<DW				0.999
Pr>DW				0.001

ID-BIG5	X			
Order	Q	Pr>Q	LM	Pr>LM
1	0.205	0.651	0.192	0.661
2	0.593	0.743	0.581	0.748
3	0.645	0.886	0.639	0.887
4	0.649	0.958	0.644	0.958
5	0.657	0.985	0.653	0.986
6	1.095	0.982	1.098	0.982
7	1.249	0.990	1.285	0.989
8	2.816	0.945	2.681	0.953
9	2.856	0.970	2.714	0.975
10	5.607	0.847	5.469	0.858
Durbin-Watson				2.091
Pr<DW				0.925
Pr>DW				0.075

ID-BIG5	log V			
Order	Q	Pr>Q	LM	Pr>LM
1	0.088	0.766	0.093	0.760
2	0.195	0.907	0.203	0.903
3	1.368	0.713	1.377	0.711
4	3.015	0.555	2.886	0.577
5	3.051	0.692	2.916	0.713
6	3.080	0.799	2.950	0.815
7	5.641	0.582	5.281	0.626
8	5.692	0.682	5.405	0.714
9	7.913	0.543	7.512	0.584
10	8.851	0.546	8.122	0.617
Durbin-Watson				2.026
Pr<DW				0.661
Pr>DW				0.339

ID-BIG5	T			
Order	Q	Pr>Q	LM	Pr>LM
1	0.001	0.982	0.001	0.970
2	0.875	0.646	0.860	0.651
3	1.019	0.797	0.998	0.802
4	2.282	0.684	2.267	0.687
5	3.053	0.692	3.059	0.691
6	3.592	0.732	3.722	0.714
7	3.933	0.788	4.078	0.771
8	5.790	0.671	5.873	0.662
9	6.063	0.734	6.030	0.737
10	8.027	0.626	7.672	0.661
Durbin-Watson				2.073
Pr<DW				0.877
Pr>DW				0.123

ID-BIG5	log Wxv			
Order	Q	Pr>Q	LM	Pr>LM
1	0.089	0.766	0.094	0.759
2	0.222	0.895	0.233	0.890
3	1.698	0.637	1.703	0.636
4	3.570	0.467	3.398	0.494
5	3.571	0.613	3.398	0.639
6	3.596	0.731	3.430	0.753
7	6.119	0.526	5.652	0.581
8	6.182	0.627	5.796	0.670
9	8.249	0.509	7.774	0.557
10	9.259	0.508	8.413	0.589
Durbin-Watson				2.027
Pr<DW				0.666
Pr>DW				0.334

ID-BIG5	Zxv			
Order	Q	Pr>Q	LM	Pr>LM
1	0.024	0.877	0.019	0.890
2	0.872	0.647	0.854	0.653
3	1.005	0.800	0.990	0.804
4	2.130	0.712	2.111	0.715
5	2.708	0.745	2.692	0.747
6	3.259	0.776	3.344	0.765
7	3.541	0.831	3.653	0.819
8	5.400	0.714	5.411	0.713
9	5.584	0.781	5.527	0.786
10	7.848	0.644	7.435	0.684
Durbin-Watson				2.075
Pr<DW				0.882
Pr>DW				0.118

ID-BIG5	log Cxv			
Order	Q	Pr>Q	LM	Pr>LM
1	0.000	0.990	0.000	0.984
2	0.028	0.986	0.026	0.987
3	0.045	0.998	0.046	0.997
4	0.571	0.966	0.587	0.965
5	1.708	0.888	1.756	0.882
6	1.787	0.938	1.819	0.936
7	3.061	0.879	3.160	0.870
8	3.061	0.931	3.161	0.924
9	4.670	0.862	4.655	0.863
10	4.712	0.910	4.683	0.911
Durbin-Watson				2.005
Pr<DW				0.529
Pr>DW				0.471

Table 3.7: DW, Q and LM tests for intraday model fits to GenBIG data.

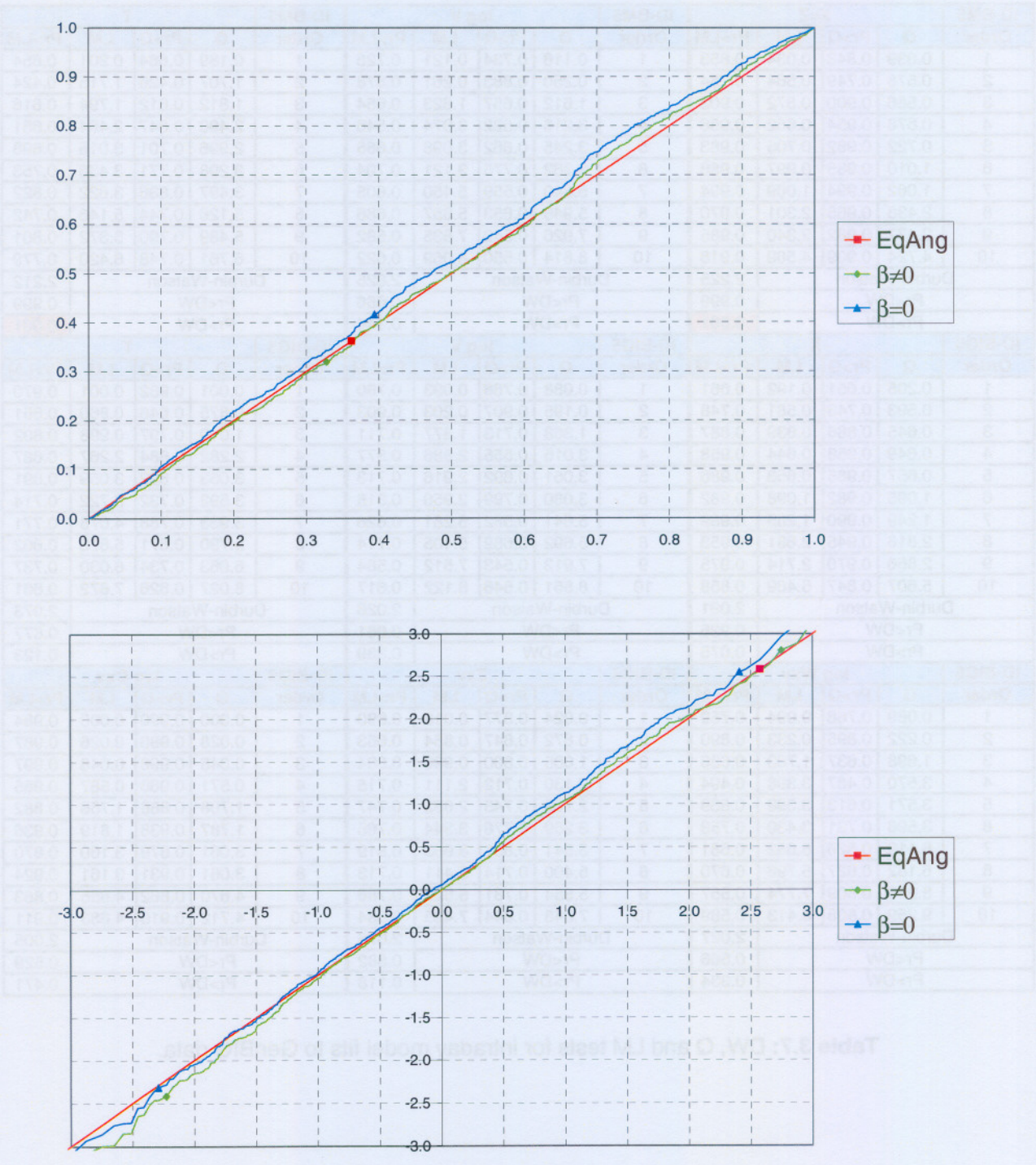

Figure 3.36: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$ ($\beta \neq 0$ and $\beta = 0$) for ID-BIG5 fit to GenBIG data.

3.3.3 IBM Data

In this section we discuss the results obtained when fitting the models of Section 3.2 to the IBM data. Sections 3.3.1 and 3.3.2 presented results of model fits to generated data and as a result we could compare the results with the true known parameter values and processes used to generate the data. This gave us an indication of the suitability of the models and accuracy of the parameter estimates. We now fit these models to real share prices. Thus, we do not possess true known parameter values and processes to compare our results to. The suitability of the models and the accuracy of the parameter estimates will be determined by considering the distributions of the residuals and the standard errors of the estimates.

ML estimates

Table 3.8 shows parameter estimates, with asymptotic standard errors in brackets, for all the model fits to the IBM data. If we take standard errors into account we see that the estimates of the D-NIG and HLC-BIG models agree reasonably well. However, as is to be expected, the standard errors of the HLC-BIG model are smaller than those of the D-NIG model. This reflects the efficiency gain obtained from using the high and low series in addition to the closing prices. We note, for example that β, ν and ϕ would not be considered significantly different from 0 according to the estimates based on the closing prices alone, but both β and ν are indeed significantly different from 0 according to the estimates based on the additional data. The estimates of the AR(1) and GARCH(1,1) parameters from the D-NOR and HLC-BM models also agree quite well with those of the D-NIG and HLC-BIG models. Again, this is in line with the consistency property of QML. However, the standard errors according to the HLC-BIG model of the two NIG parameters β and ψ suggest that these are indeed required and that we may expect the NIG distribution based models to fit better than the normal distribution based models.

Turning to the intraday models, although the parameter estimates obtained by most of the model fits agree reasonably well with each other, there are a few exceptions. The estimates for β and α_0 obtained by the ID-BM model fits are considerably lower than the corresponding estimates obtained by the other model fits. Furthermore, it is again clear that the standard errors generally get smaller as more and more data is used in the models. For instance, the standard error for α_1 shrinks from 0.0185 for the D-NIG model using only daily returns, to 0.0119 for the HLC-BIG model using daily returns as well as high and low returns, to 0.0088 for the ID-BIG model using 5 minute intraday returns and realised volatilities. Thus, using the intraday data in addition to the total returns data pays off in terms of substantially more accurate parameter estimates. Comparing the standard errors for the 5, 10 and 30 minute data, there seems to be a tendency for smaller values

accompanying shorter periods (larger I) but most of the differences are slight. These estimates (and their standard errors) compare quite well with those obtained by using the HLC-BIG model. This suggests that for the purpose of estimation of the parameters of the ID-BIG model it is not critically important to use a large number of intraday periods and that one may even get by with only using high, low and close data rather than full high frequency data. Further, the consistency of the different sets of estimates suggests that the ID-BIG model with the present specifications of $\hat{\mu}_n$ and $\hat{h}_n$ fits this data well.

Note that the estimate of β is negative. This is in line with the results of Engle and Ng (1993) and conforms to the general finding that the news impact curve is asymmetric with negative innovations having greater volatility impacts than positive ones. The estimates of ϕ suggest that the autoregressive part of the AR(1) component is not needed and one can simplify and refit the models with $\phi = 0$ to begin with. The resulting estimates, of which an extraction is given in Table 3.9, agree quite closely with their counterparts in Table 3.8 and since no new matters arise we omit further details.

Quality of fit

Figure 3.38 shows the PIT and Q-Q plots of the daily return residuals corresponding to the D-NOR and HLC-BM model fits, while Figures 3.38 and 3.39 show the PIT and Q-Q plots for the residuals and estimates of the D-NIG and HLC-BIG model fits respectively. It is clear that the distributions of the residuals and estimates obtained from the D-NIG and HLC-BIG models (note that the plots for $\hat{X}_n$ and $\hat{Z}_{X,n}$ are nearly identical) are closer to their model implied distributions than those obtained from the D-NOR and HLC-BM models. This means that the models based on NIG-distributed innovations (or underlying BIG processes) lead to residuals and estimates that agree better with their model assumptions than models based on normally distributed innovations (or underlying BM processes).

Figures 3.40 to 3.42 show that the distributions of $\hat{V}_n$ (and to a lesser degree $\hat{T}_n$) obtained from fitting the ID-BM model to the 5, 10 and 30 minute IBM data differ greatly from their model implied distribution, so that the ID-BM model is clearly inadequate. Thus within the ID-BIG model the choice $\psi = \infty$ must be rejected for the IBM data.

Figures 3.43 to 3.48 show the PIT and Q-Q plots to check the fit of the ID-BIG model to the 5, 10 and 30 minute IBM data by means of the $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ residuals and the latent variable estimates $\hat{W}_{XV,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$. In all cases the model fit seems adequate.

Parameter	D-NOR	D-NIG
β	-0.3914 (0.5040)	-0.1627 (0.1910)
ψ	-	2.1563 (0.7251)
ν	0.0071 (0.0101)	0.0025 (0.0037)
ϕ	-0.0253 (0.0399)	-0.0254 (0.0331)
$\alpha_0 x 10^5$	0.7695 (0.7203)	0.7953 (0.4386)
α_1	0.0553 (0.0270)	0.0496 (0.0185)
β_1	0.9203 (0.0319)	0.9311 (0.0250)
Parameter	**HLC-BM**	**HLC-BIG**
β	-0.5670 (0.0929)	-0.2241 (0.0684)
ψ	-	1.7424 (0.0933)
ν	0.0096 (0.0016)	0.0036 (0.0012)
ϕ	-0.0326 (0.0337)	-0.0268 (0.0302)
$\alpha_0 x 10^5$	0.8115 (0.3089)	1.3346 (0.3999)
α_1	0.0577 (0.0095)	0.0802 (0.0119)
β_1	0.8931 (0.0177)	0.8798 (0.0195)
Parameter	**ID-BM30**	**ID-BIG30**
β	-0.6167 (0.0861)	-0.2175 (0.0601)
ψ	∞	1.5821 (0.0786)
ν	0.0113 (0.0016)	0.0035 (0.0010)
ϕ	-0.0065 (0.0430)	-0.0346 (0.0290)
$\alpha_0 x 10^5$	0.4746 (0.3892)	1.7040 (0.5372)
α_1	0.0717 (0.0135)	0.0964 (0.0149)
β_1	0.8923 (0.0215)	0.8577 (0.0252)
Parameter	**ID-BM10**	**ID-BIG10**
β	-0.5999 (0.0748)	-0.2552 (0.0514)
ψ	∞	1.6871 (0.0698)
ν	0.0110 (0.0014)	0.0042 (0.0008)
ϕ	0.0379 (0.0584)	-0.0272 (0.0273)
$\alpha_0 x 10^5$	0.1798 (0.2730)	1.1125 (0.2843)
α_1	0.0737 (0.0106)	0.0960 (0.0102)
β_1	0.8997 (0.0151)	0.8737 (0.0144)
Parameter	**ID-BM5**	**ID-BIG5**
β	-0.5815 (0.0696)	-0.2912 (0.0490)
ψ	∞	1.8118 (0.0728)
ν	0.0108 (0.0013)	0.0050 (0.0007)
ϕ	-0.0077 (0.0540)	-0.0280 (0.0268)
$\alpha_0 x 10^5$	0.2737 (0.2373)	0.9349 (0.2511)
α_1	0.0738 (0.0092)	0.0950 (0.0088)
β_1	0.9001 (0.0129)	0.8807 (0.0119)

Table 3.8: ML Estimates of all model fits to IBM data.

Parameter	D-NIG	HLC-BIG
β	-0.1542 (0.1892)	-0.2187 (0.0694)
ψ	2.1422 (0.7188)	1.7426 (0.0936)
ν	0.0024 (0.0037)	0.0035 (0.0012)
ϕ	0.0000 -	0.0000 -
$\alpha_0 x 10^5$	0.7994 (0.4611)	1.3628 (0.3817)
α_1	0.0504 (0.0193)	0.0802 (0.0117)
β_1	0.9304 (0.0264)	0.8792 (0.0188)
Parameter	**ID-BM5**	**ID-BIG5**
β	-0.5839 (0.0703)	-0.2873 (0.0497)
ψ	∞	1.8109 (0.0728)
ν	0.0108 (0.0013)	0.0050 (0.0007)
ϕ	0.0000 -	0.0000 -
$\alpha_0 x 10^5$	0.2706 (0.2382)	0.9656 (0.2612)
α_1	0.0737 (0.0092)	0.0948 (0.0089)
β_1	0.9002 (0.0129)	0.8802 (0.0122)

Table 3.9: ML Estimates for model fits to IBM data when $\phi = 0$.

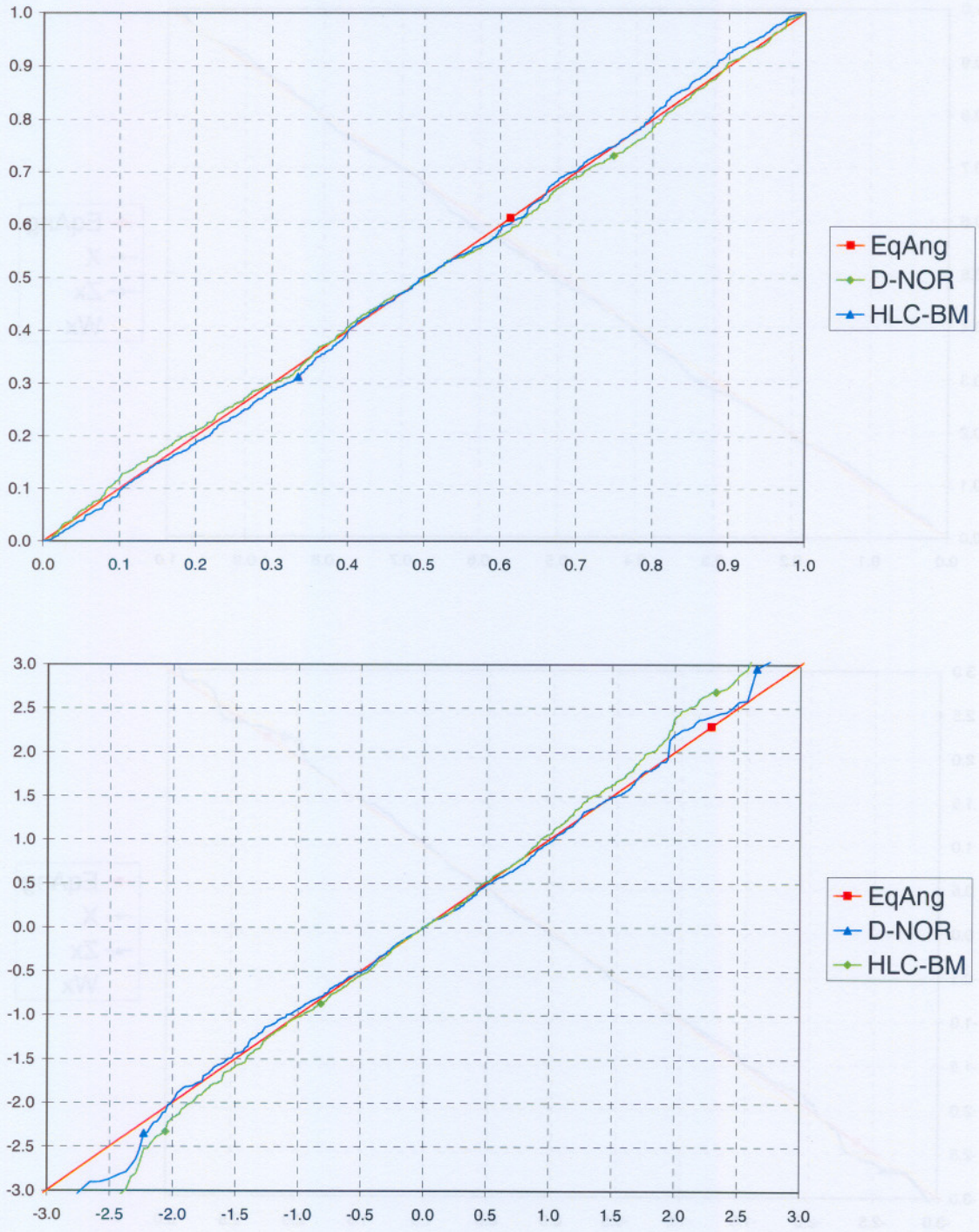

Figure 3.37: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n \equiv \hat{Z}_n$ residuals checking D-NOR and HLC-BM fits to IBM data.

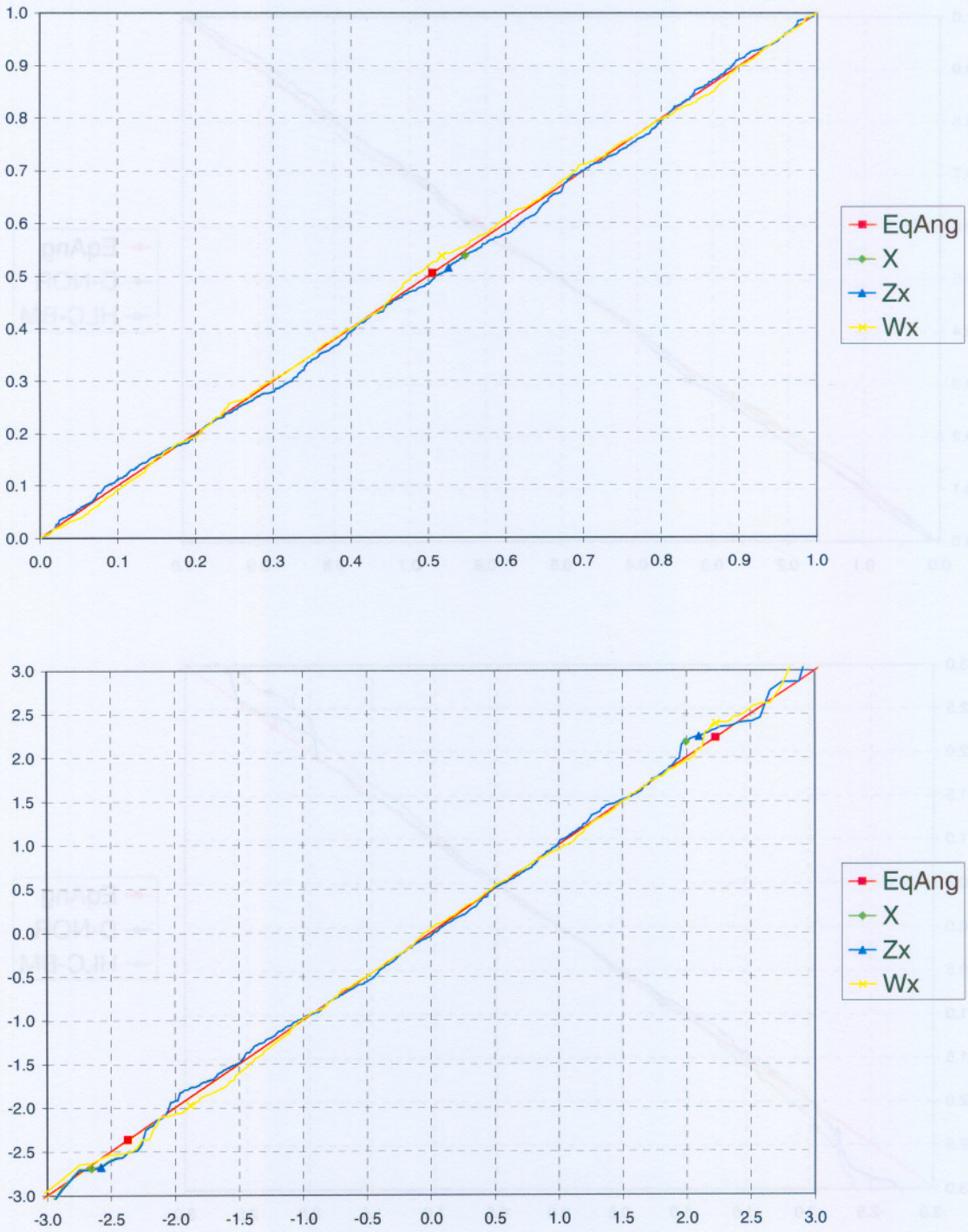

Figure 3.38: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_{X,n}$ and $\hat{W}_{X,n}$ checking D-NIG fit to IBM data.

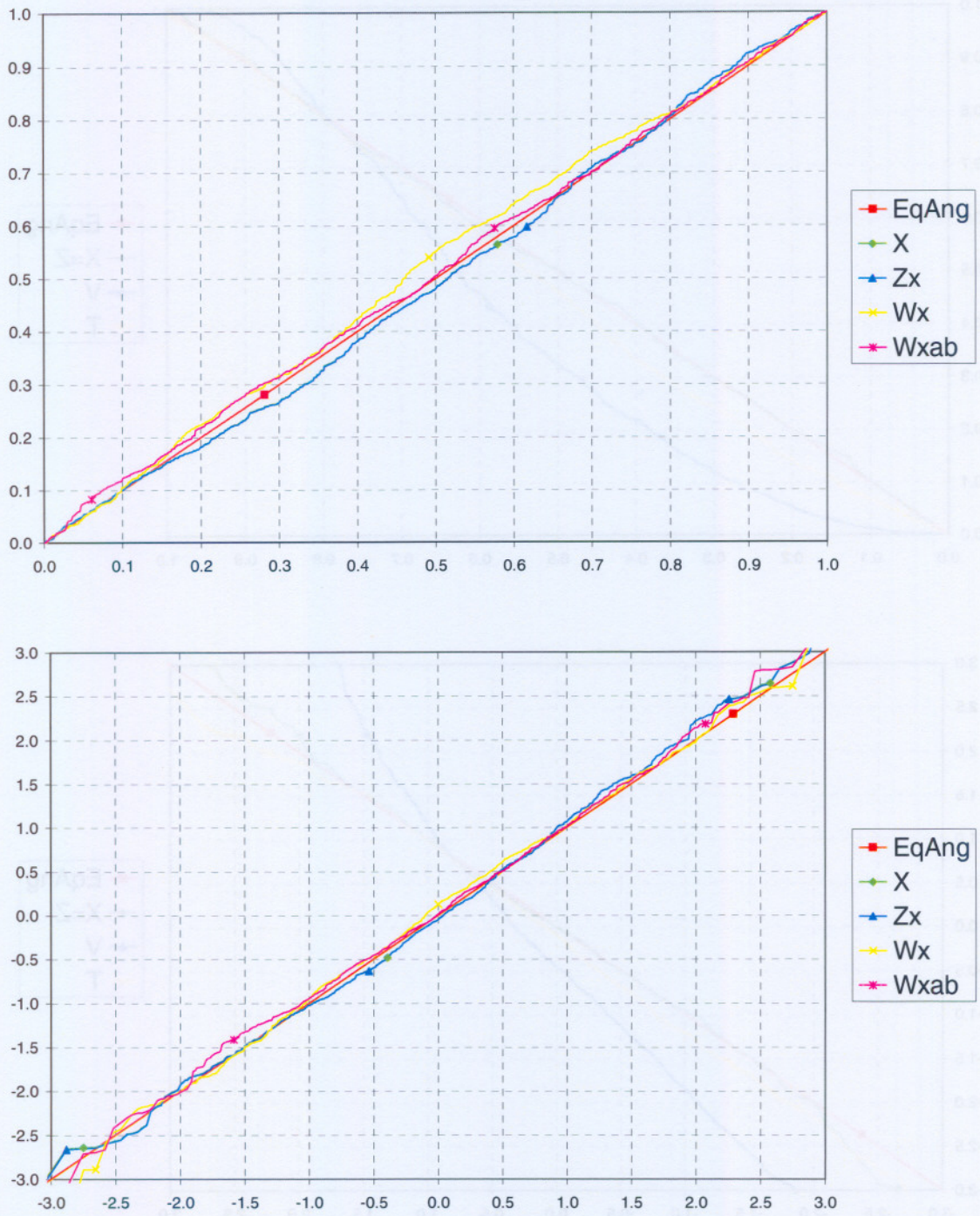

Figure 3.39: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_{X,n}$, $\hat{W}_{X,n}$ and $\hat{W}_{XAB,n}$ checking HLC-BIG fit to IBM data.

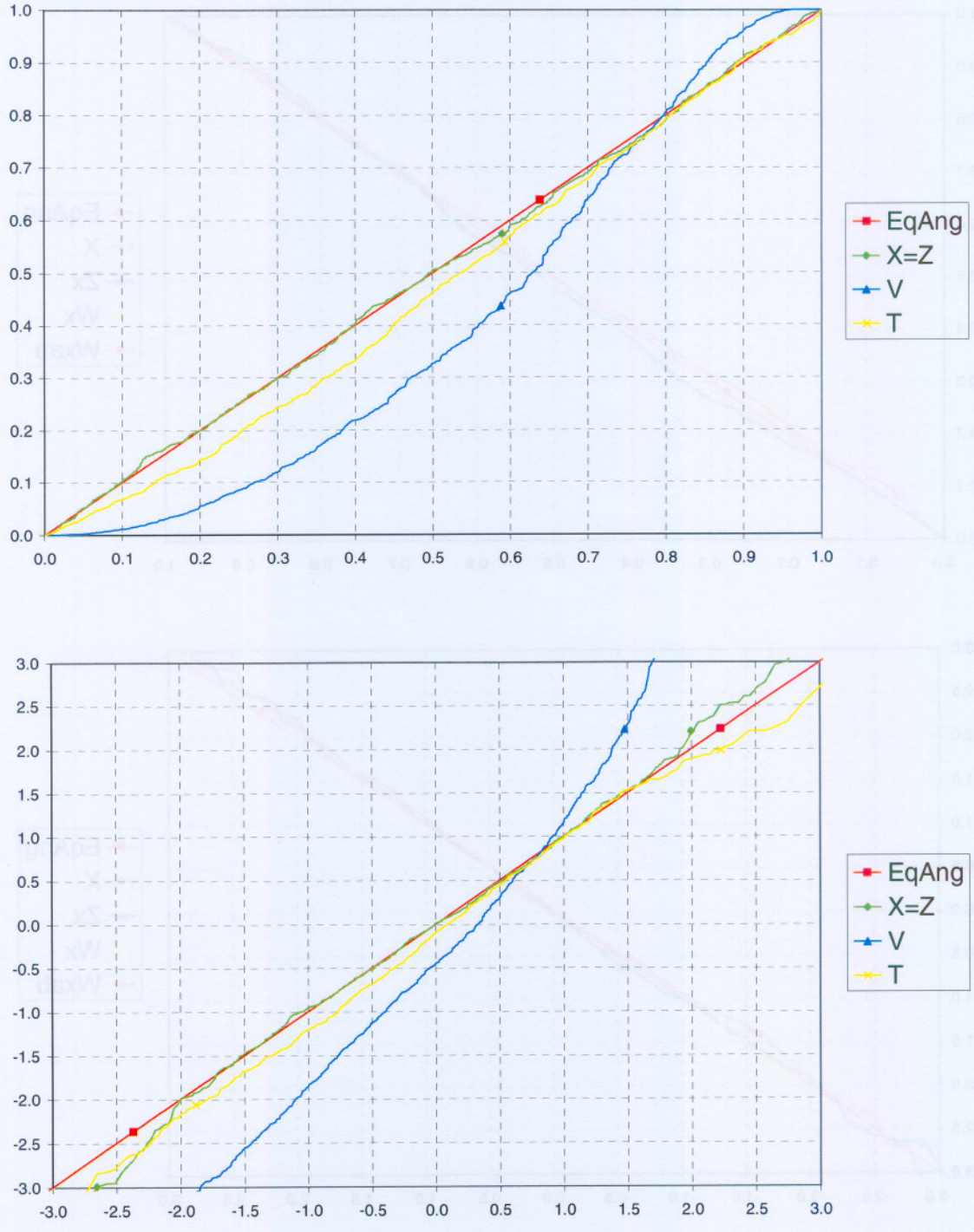

Figure 3.40: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BM30 fit to IBM data.

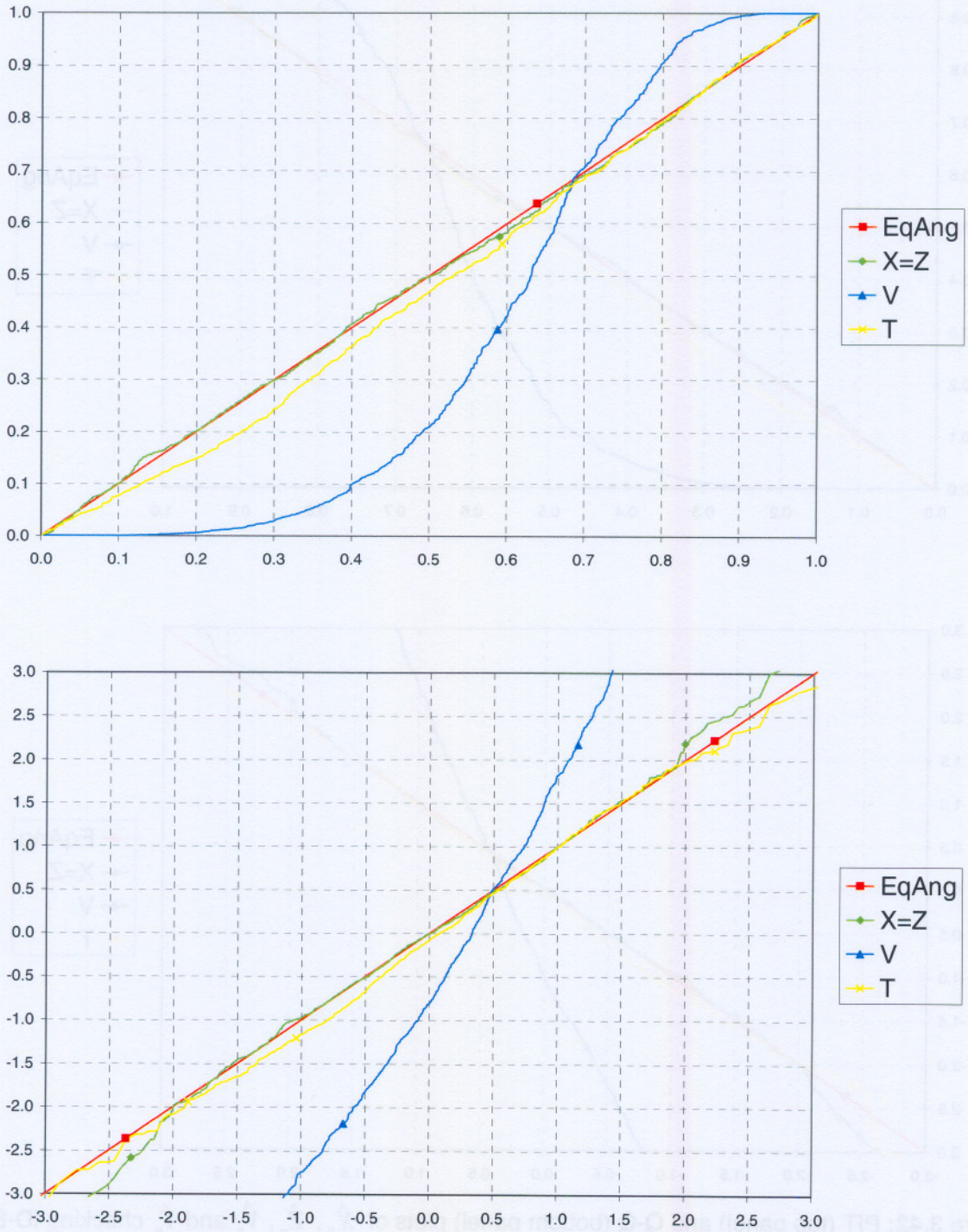

Figure 3.41: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BM10 fit to IBM data.

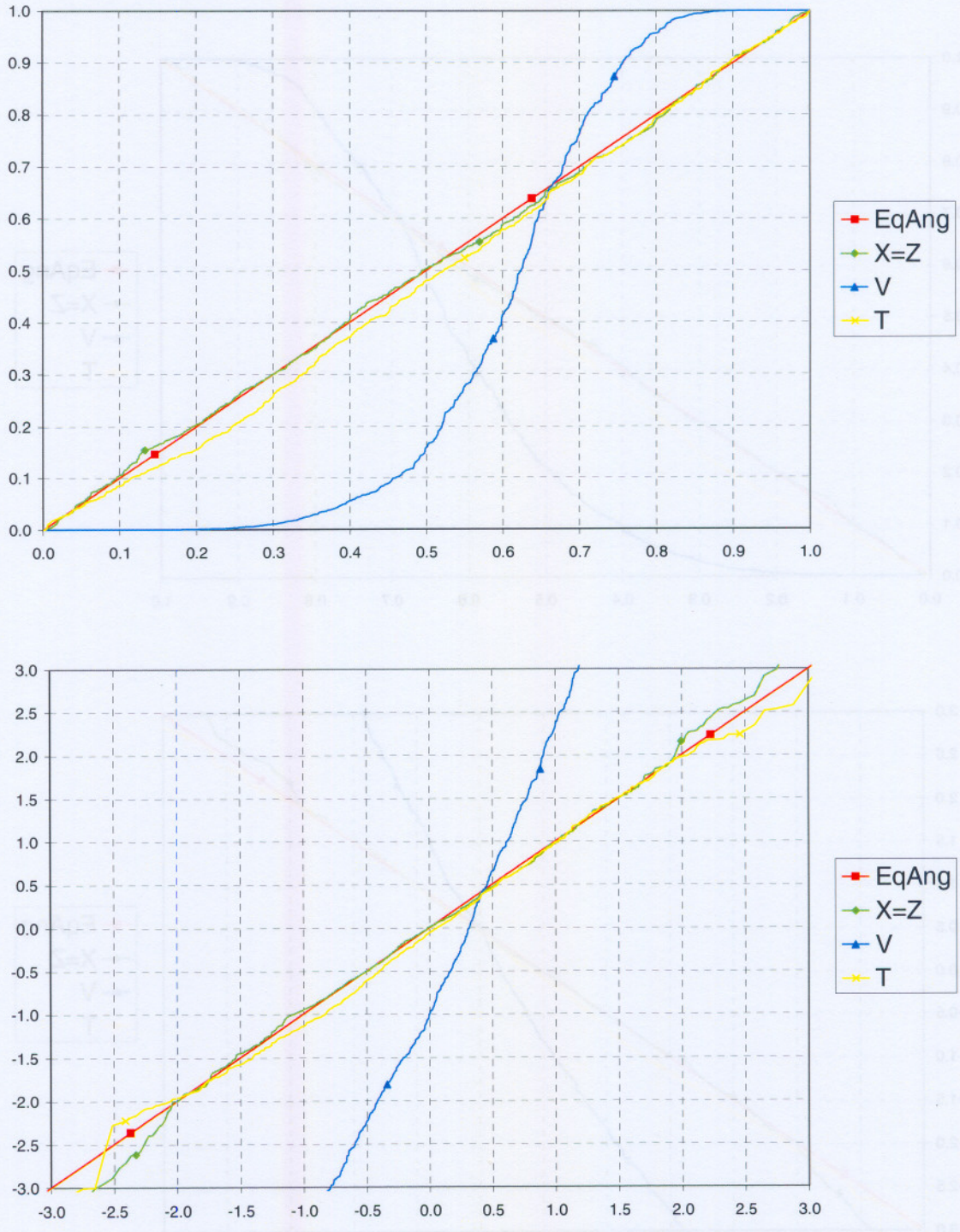

Figure 3.42: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{Z}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BM5 fit to IBM data.

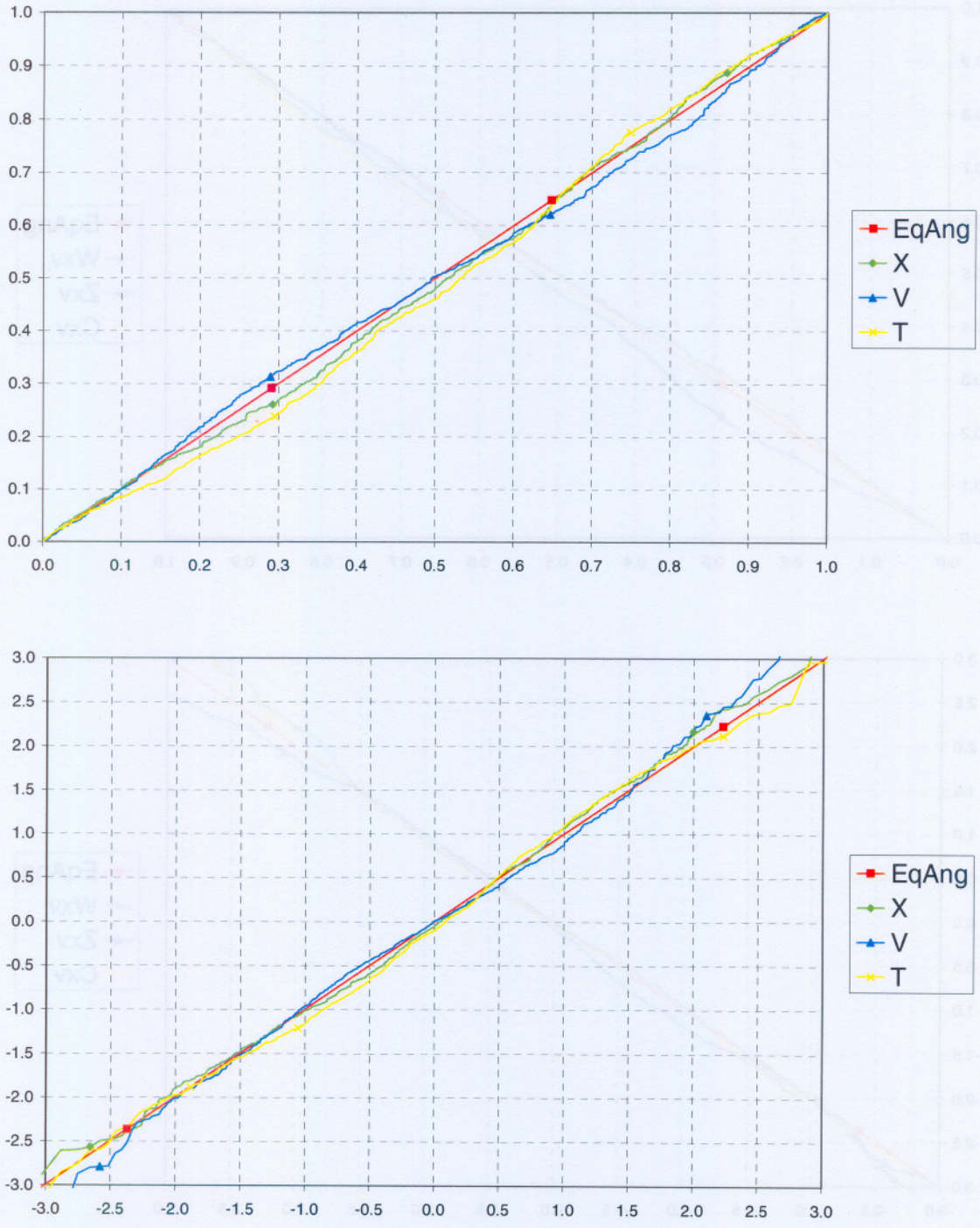

Figure 3.43: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BIG30 fit to IBM data.

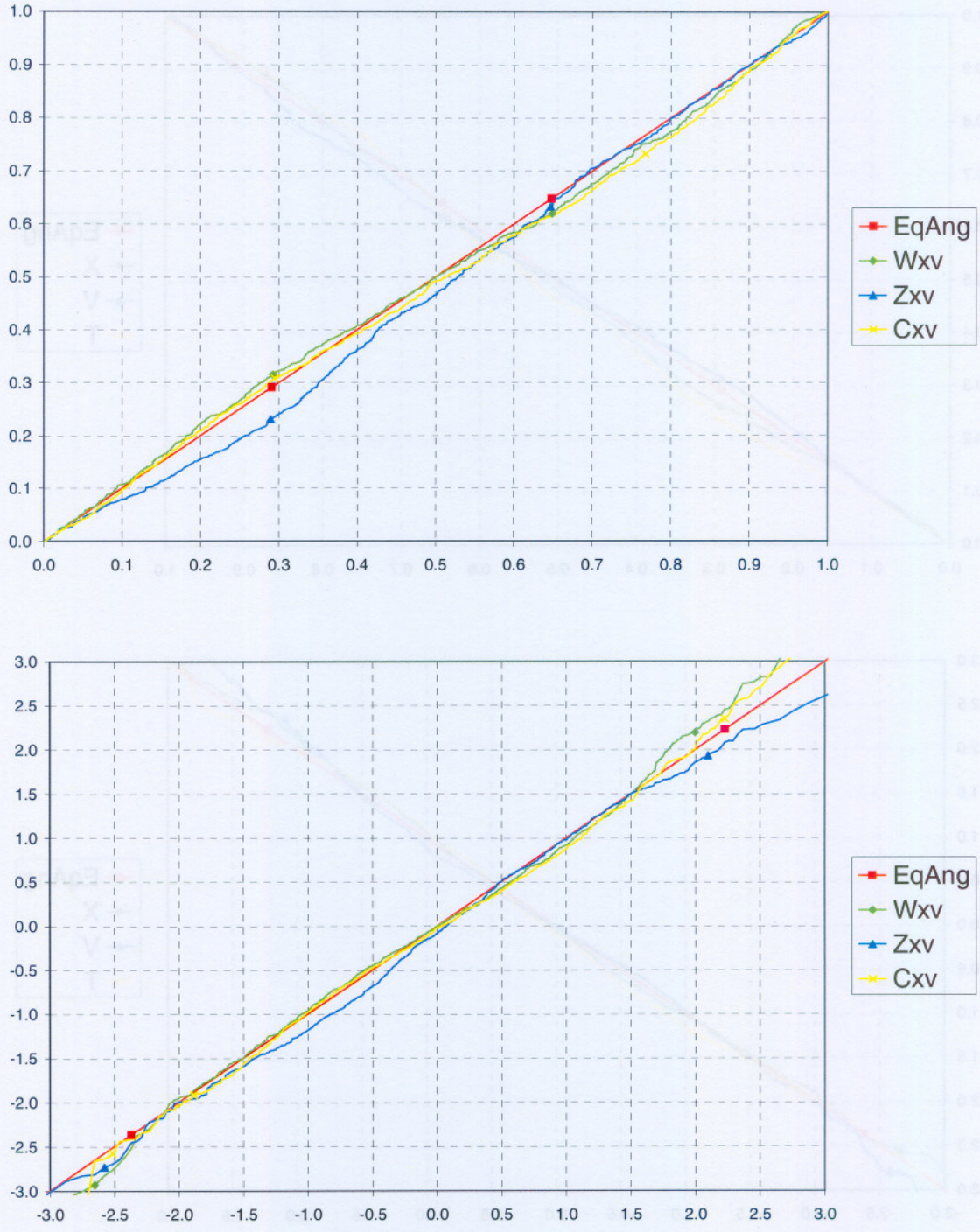

Figure 3.44: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{XV,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking ID-BIG30 fit to IBM data.

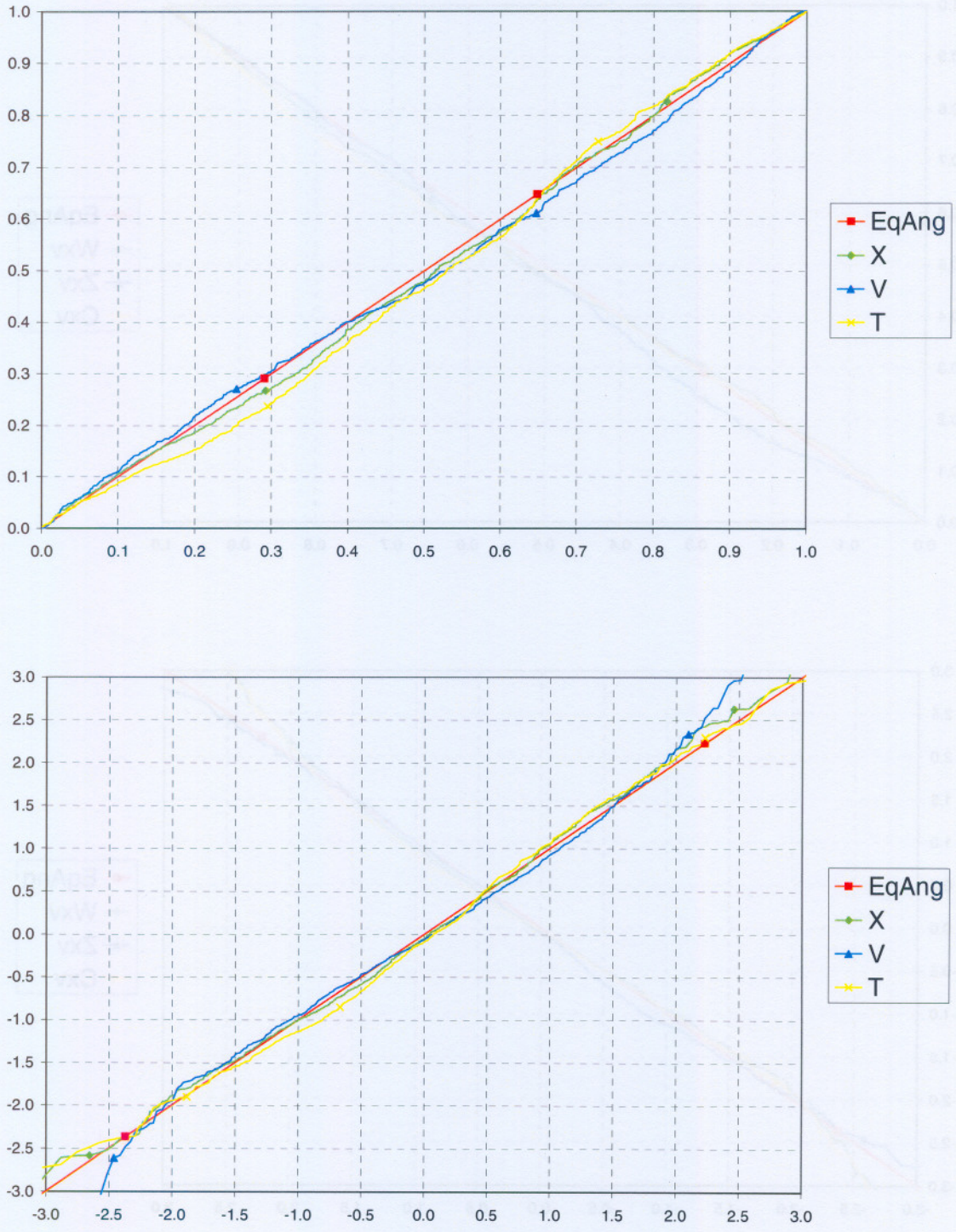

Figure 3.45: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BIG10 fit to IBM data.

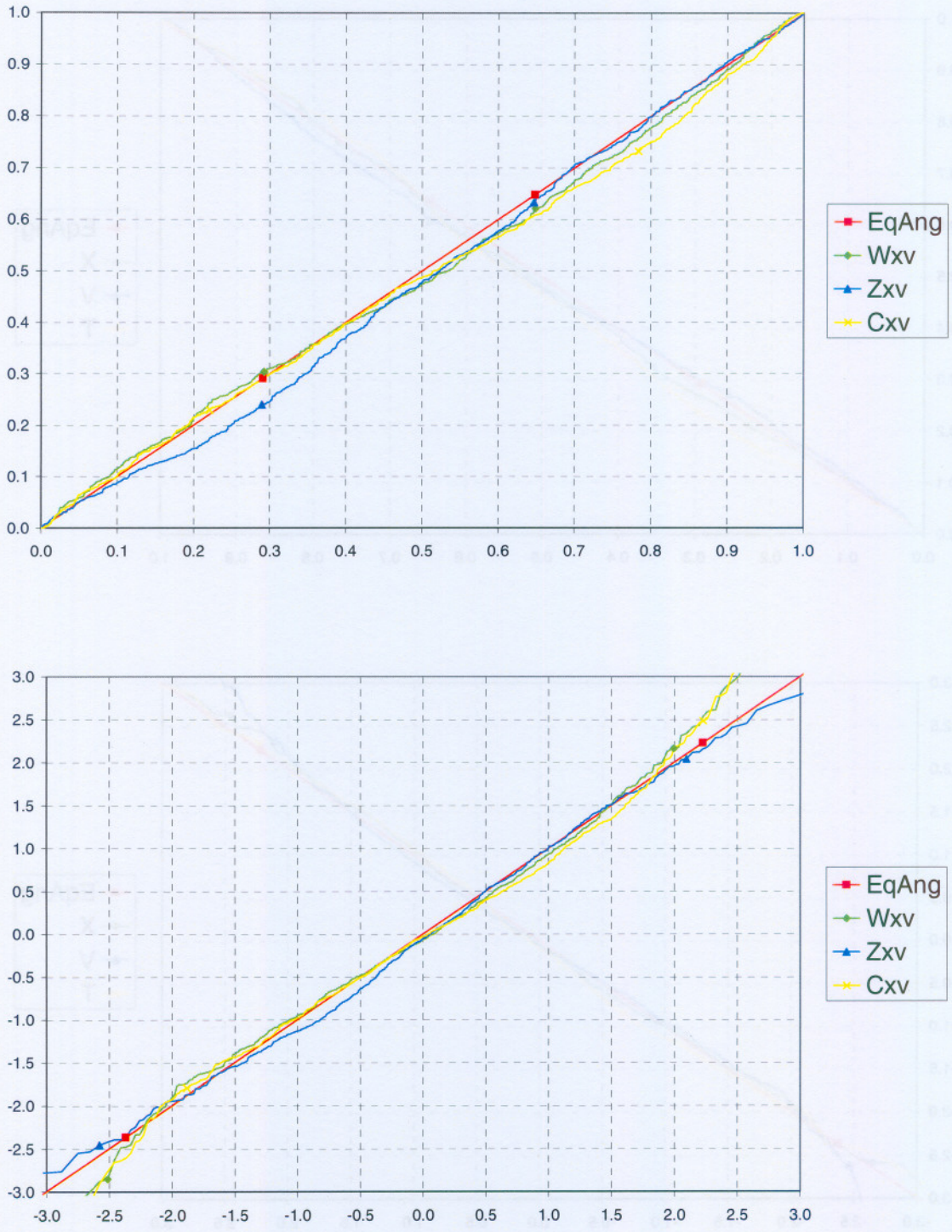

Figure 3.46: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{XV,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking ID-BIG10 fit to IBM data.

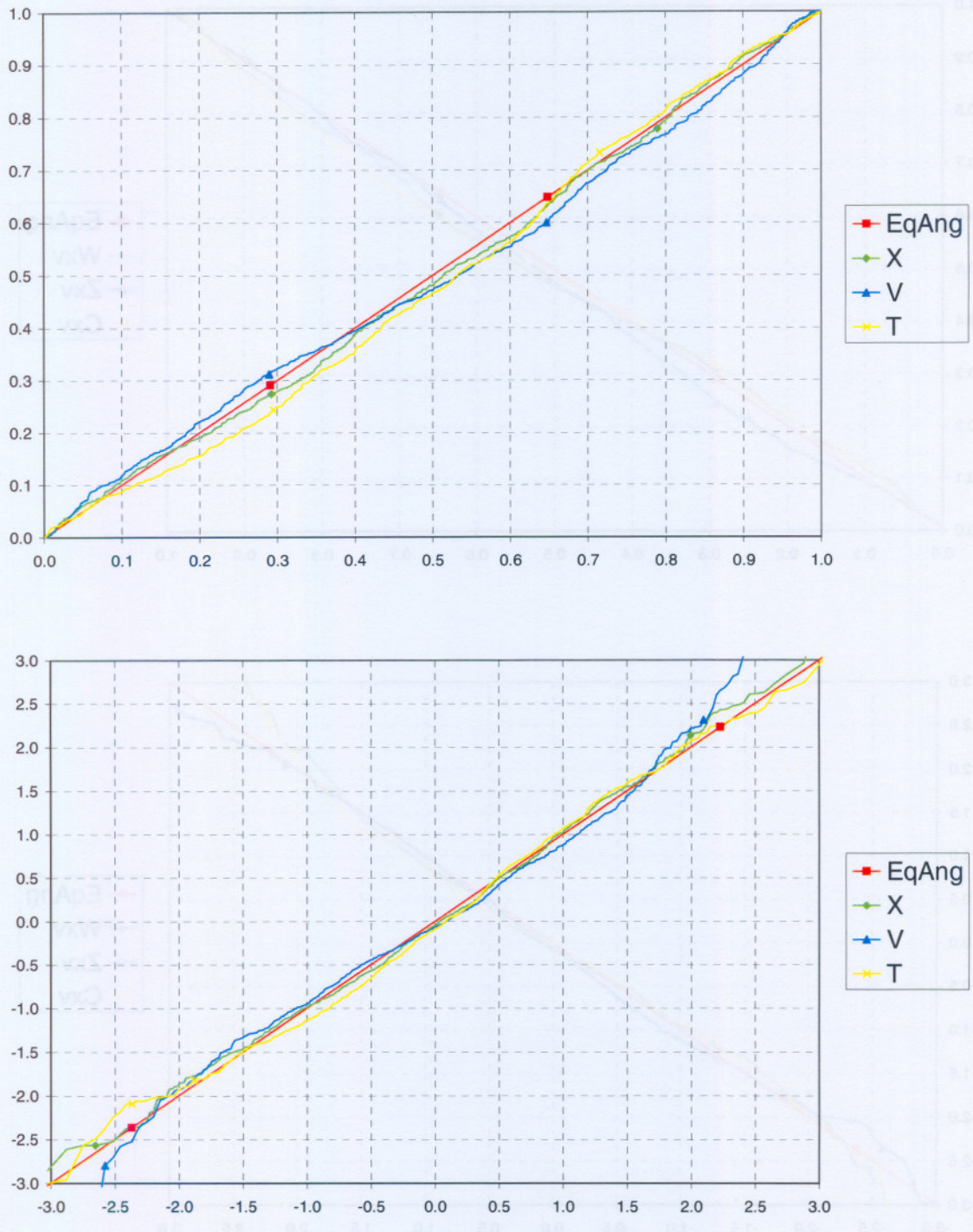

Figure 3.47: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking ID-BIG5 fit to IBM data.

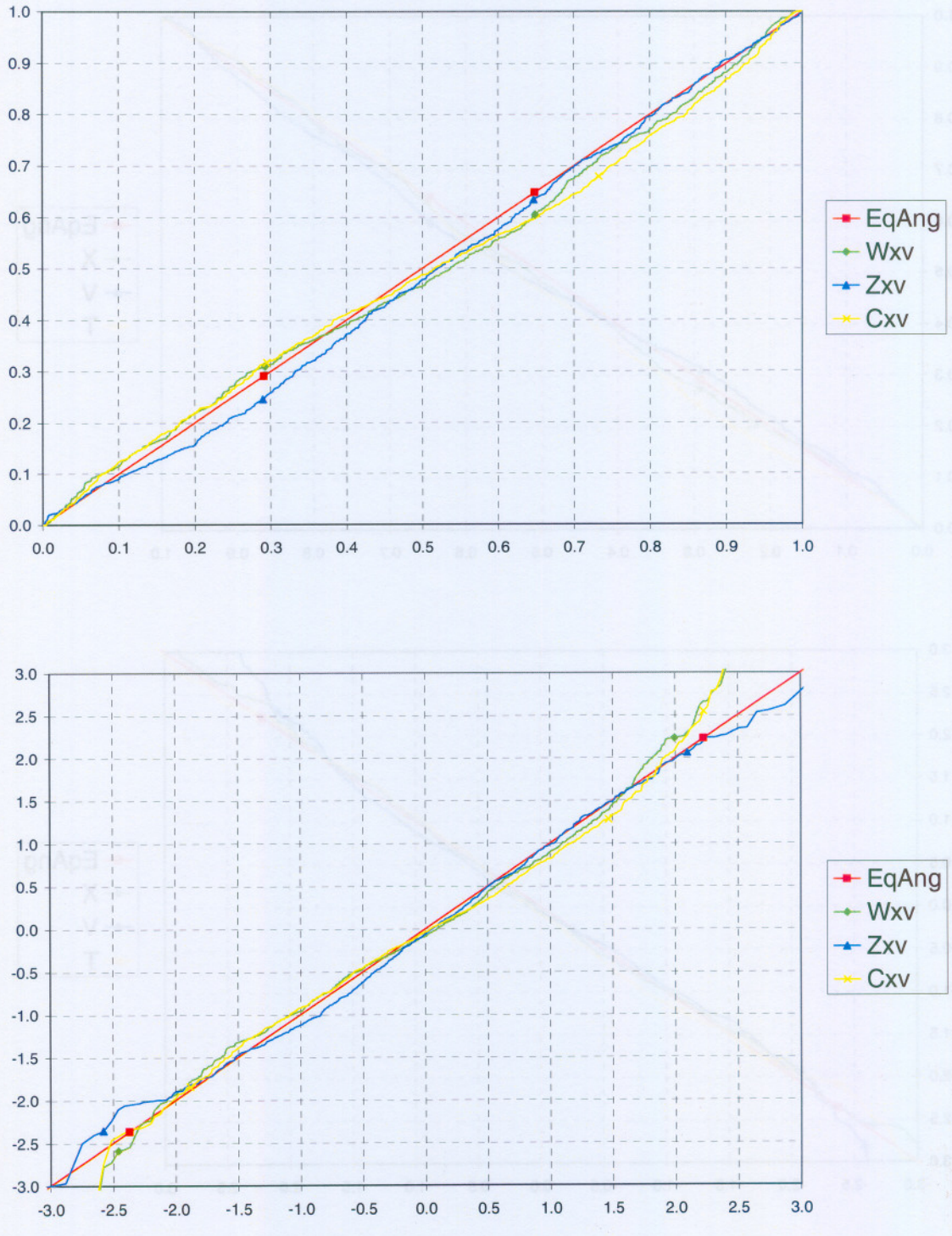

Figure 3.48: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{XV,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking ID-BIG5 fit to IBM data.

Volatility estimates

We also computed the estimated volatility adjustment factors $\hat{W}_{X,n}$, $\hat{W}_{XAB,n}$ (with $\mu_n = 0$) and $\hat{W}_{XV,n}$, as well as the estimates $\hat{Z}_{X,n}$, $\hat{C}_{X,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$, for the D-NIG, HLC-BIG and ID-BIG model fits to the IBM data and substituted the relevant parameter estimates and data values into the h_n specification to obtain the $\hat{h}_n$'s. Remember that we do not possess true known parameter values and processes to compare our results to, so that we cannot, for instance, compare the estimated $\hat{W}_{XV,n}$'s to true W_n's. However, we may still compare the volatility estimates with each other to observe the influence of using higher-frequency data in our models.

Figure 3.49 compares $\hat{h}_n$ with $\hat{h}_n\hat{W}_{X,n}$ obtained from the D-NIG model fit on a log-scale over a period of the series when volatility varied substantially. The top panel of Figure 3.49 shows that the adjusted volatility estimator $\hat{h}_n\hat{W}_{X,n}$ tends to lead the ordinary GARCH volatility estimator $\hat{h}_n$ by one day, diminishing the one-day lag that is characteristic of GARCH volatility estimates. However, apart from diminishing this one-day lag, $\hat{h}_n\hat{W}_{X,n}$ does not adjust $\hat{h}_n$ very much and, as can be seen from the bottom panel of Figure 3.49, it is still not close to the daily realised volatility R_n (calculated from 5 minute returns). We expect the adjusted volatility estimator to be closer to the realised volatility when fitting models that use more data, for instance the HLC-BIG model. Figure 3.50 compares $\hat{h}_n\hat{W}_{X,n}$ and $\hat{h}_n\hat{W}_{XAB,n}$, obtained from the HLC-BIG fit, with the 5 minute realised volatility R_n over the same time period as Figure 3.49, also on a log-scale . The top panel of Figure 3.50 shows the effect that the adjustment factors $\hat{W}_{X,n}$ and $\hat{W}_{XAB,n}$ have on $\hat{h}_n$. $\hat{h}_n\hat{W}_{X,n}$, which uses only close prices, still does not adjust $\hat{h}_n$ very much, except for diminishing the one-day lag. However, $\hat{h}_n\hat{W}_{XAB,n}$, which incorporates the high and low prices also, adjusts $\hat{h}_n$ considerably in addition to diminishing the one-day lag. The bottom panel of Figure 3.50 shows that $\hat{h}_n\hat{W}_{XAB,n}$ is considerably closer than $\hat{h}_n\hat{W}_{X,n}$ to R_n. This can also be seen in Figure 3.51, which plots $\hat{h}_n\hat{W}_{X,n}$ and $\hat{h}_n\hat{W}_{XAB,n}$ respectively against R_n. Also note from Figure 3.39 above that the PIT and Q-Q plots for $\hat{W}_{XAB,n}$ are closer to the equi-angular line than the plots for $\hat{W}_{X,n}$, so that the distribution of $\hat{W}_{XAB,n}$ are closer to its model implied distribution.

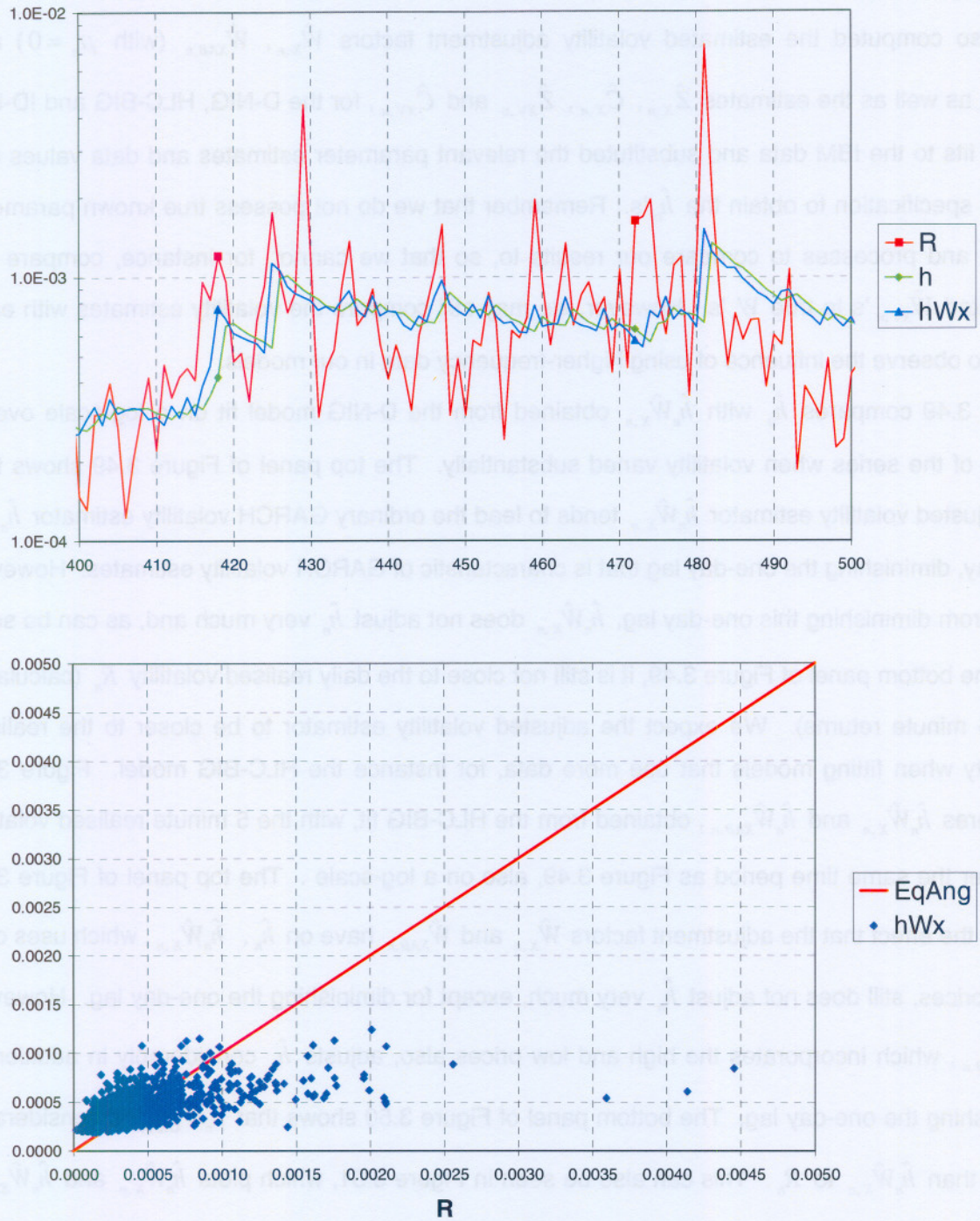

Figure 3.49: Plots of $\hat{h}_n$ and $\hat{h}_n \hat{W}_{X,n}$ (top panel) and $\hat{h}_n \hat{W}_{X,n}$ (bottom panel) against R_n for D-NIG fit to IBM data.

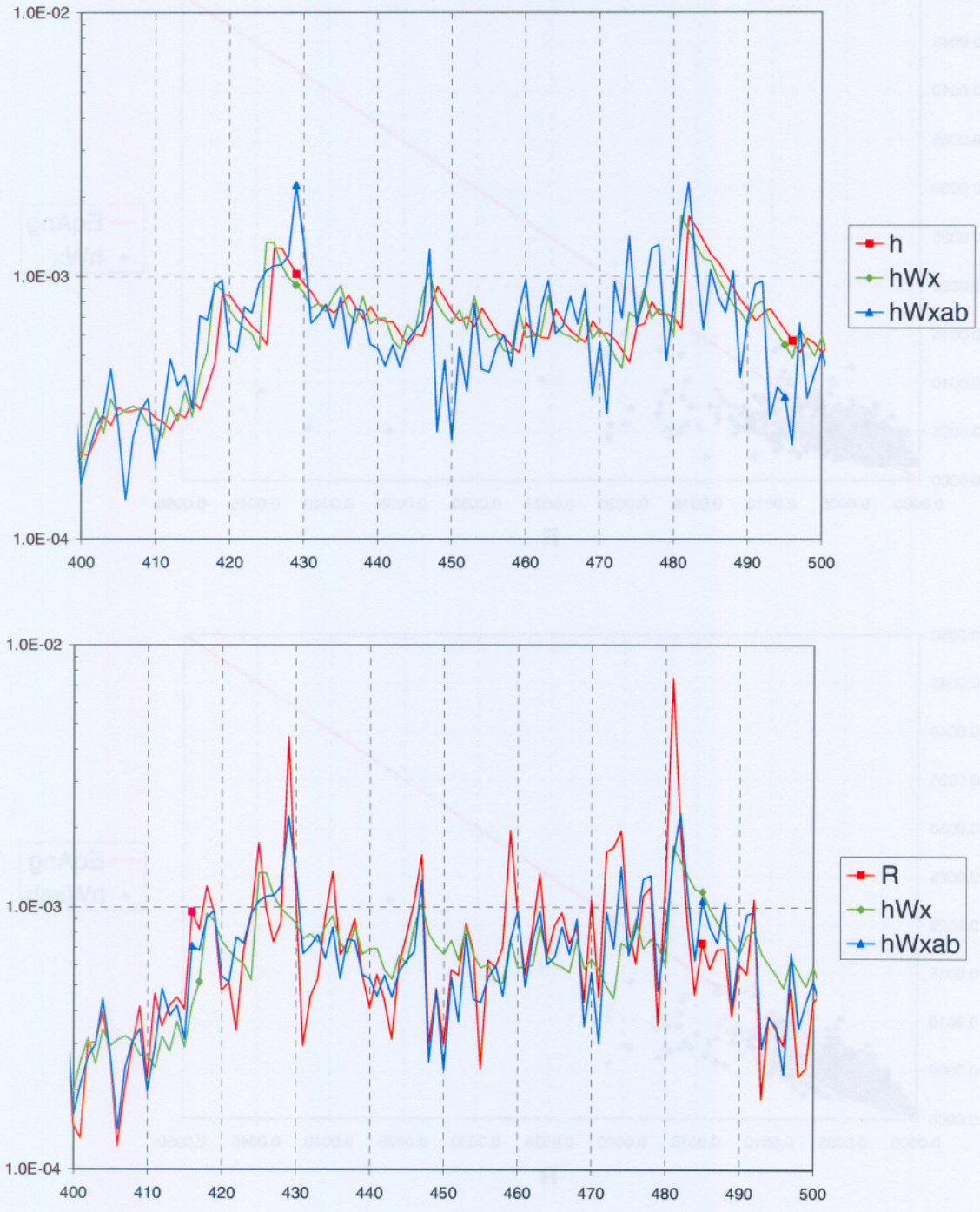

Figure 3.50: Plots of $\hat{h}_n$, $\hat{h}_n \hat{W}_{X,n}$ and $\hat{h}_n \hat{W}_{XAB,n}$ (top panel) and R_n, $\hat{h}_n \hat{W}_{X,n}$ and $\hat{h}_n \hat{W}_{XAB,n}$ (bottom panel) for HLC-BIG fit to IBM data.

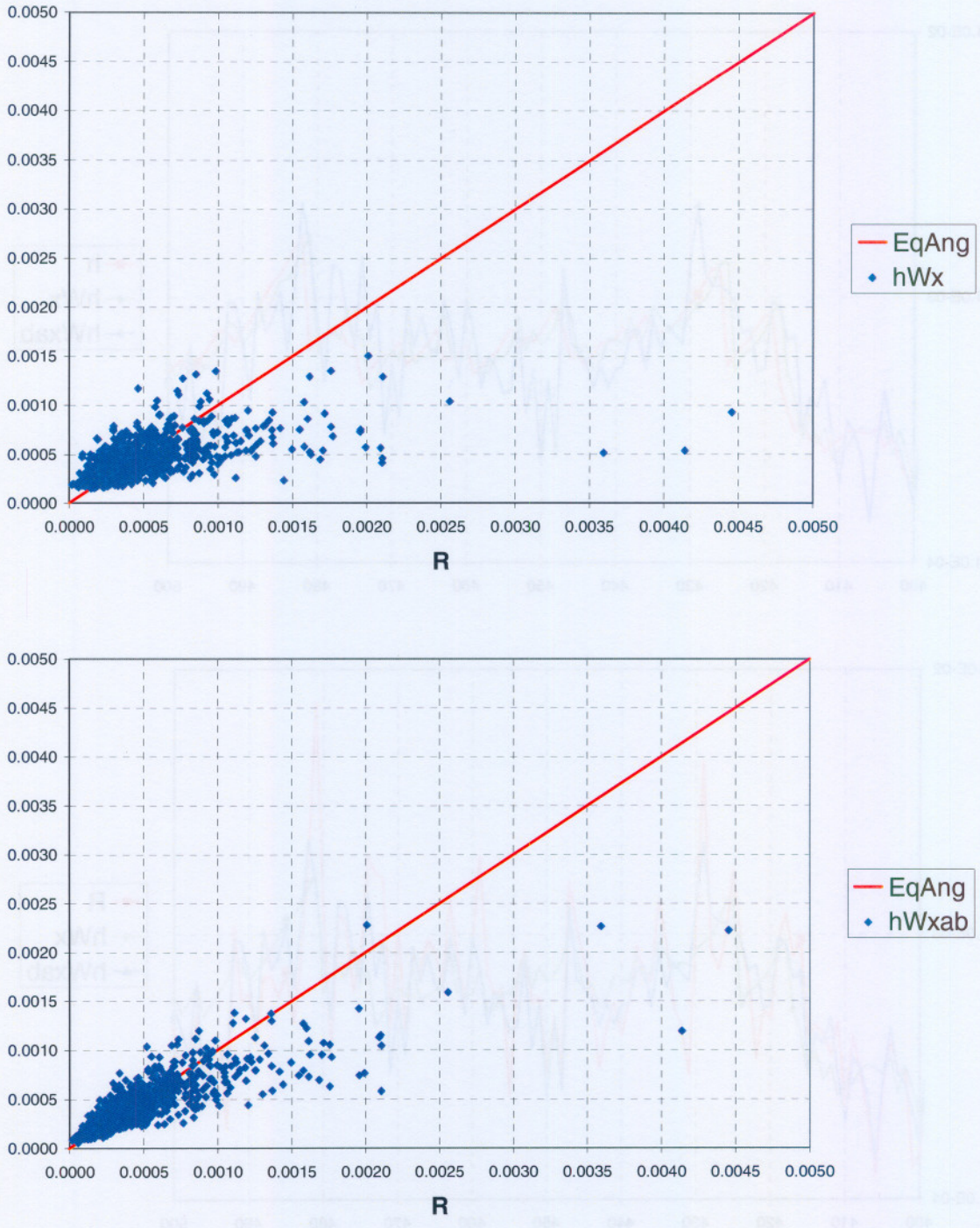

Figure 3.51: Plots of $\hat{h}_n \hat{W}_{X,n}$ (top panel) and $\hat{h}_n \hat{W}_{XAB,n}$ (bottom panel) against R_n for HLC-BIG fit to IBM data.

Turning to the ID-BIG model fit to the IBM data, consider Figure 3.52, which shows the estimated expected volatility $\hat{h}_n$, the estimated actual volatility $\hat{h}_n\hat{W}_{XV,n}$ (both based on the ID-BIG model fit to 5 minute data) and the realised volatility R_n over days 400 to 500. Again, the close agreement between $\hat{h}_n\hat{W}_{XV,n}$ and R_n is clearly visible. Comparing Figure 3.52 with the bottom panel of Figure 3.50 above we see that $\hat{h}_n\hat{W}_{XV,n}$ tracks R_n much better than $\hat{h}_n\hat{W}_{X,n}$ and $\hat{h}_n\hat{W}_{XAB,n}$ obtained from the HLC-BIG model fit, again indicating the increase in accuracy as more data becomes available. The role of $\hat{h}_n$ is also clear, namely it represents a smoothed version of the more highly variable daily actual volatility. Figures 3.53 to 3.55 show scatter plots of the 5, 10 and 30 minute realised volatilities and the corresponding ID-BIG model estimates $\hat{h}_n\hat{W}_{XV,n}$. Comparing the points with the equi-angular line, the agreement becomes better for shorter periods and is excellent for the 5 minute data, barring only the four largest values. This illustrates the arguments at the end of Section 3.2.3 and confirms that the realised volatility based on short intraday periods is indeed a good proxy for the actual volatility, at least in this example. Realised volatilities are simple to compute and require very little by way of model assumptions and analysis compared to that of the ID-BIG model presented here. Clearly, realised volatility is an important tool for volatility analysis when high-frequency data is available.

As mentioned earlier, since in this section we are not in the simulated data situation of Sections 3.3.1 and 3.3.2, we cannot compare the volatility estimates with their true values, which are unknown. We can, however, compare the results for different intraday data frequencies to get an indication of the stability of these volatility estimates. Since the highest-frequent intraday data we have available is 5 minute data, we will compare the results of the ID-BIG model fit to the 5 minute data to the results of the ID-BIG model fit to the 10 minute and 30 minute data.

Figures 3.56 to 3.59 show scatter plots of the realised volatilities R_n, the ordinary GARCH volatility estimates $\hat{h}_n$, the estimates of the random impact factors $\hat{W}_{XV,n}$ and the actual volatility estimates $\hat{h}_n\hat{W}_{XV,n}$ respectively, comparing the results of fits to the 10 and 30 minute period data with those of the fit to the 5 minute period data. In all cases, the correlation between the values for the 10 minute data and the 5 minute data is greater than the correlation between the values for the 30 minute data and the 5 minute data. This illustrates that R_n, $\hat{h}_n$, $\hat{W}_{XV,n}$ and $\hat{h}_n\hat{W}_{XV,n}$ become more stable as the intraday period shortens.

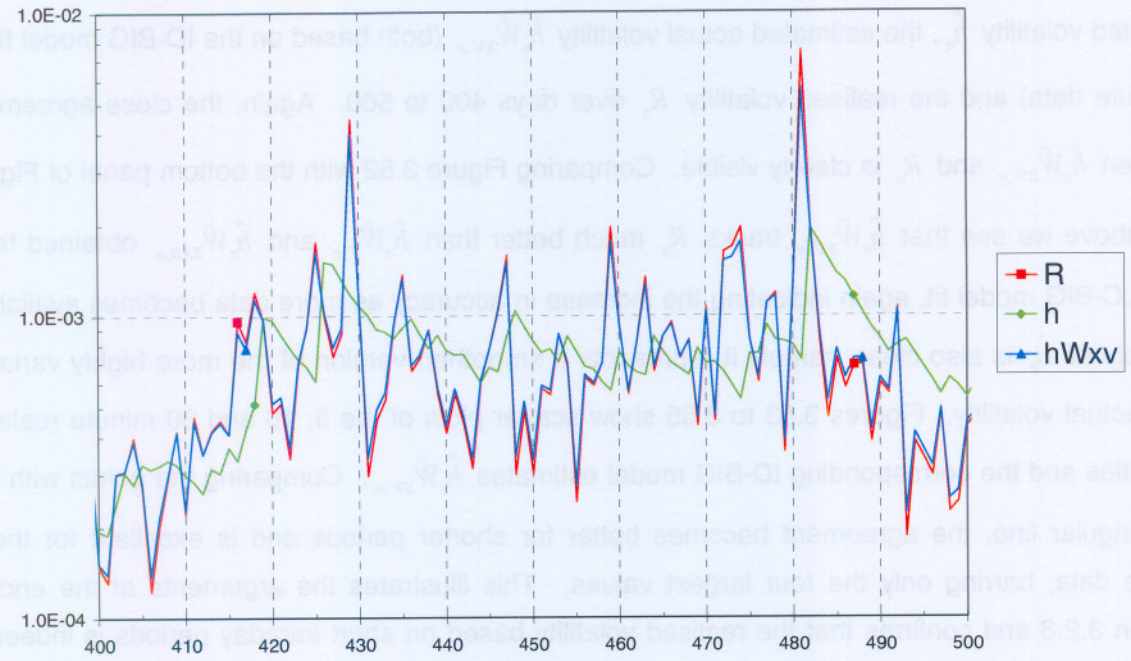

Figure 3.52: Comparing $\hat{h}_n$, $\hat{h}_n \hat{W}_{XV,n}$ and R_n for ID-BIG5 fit to IBM data.

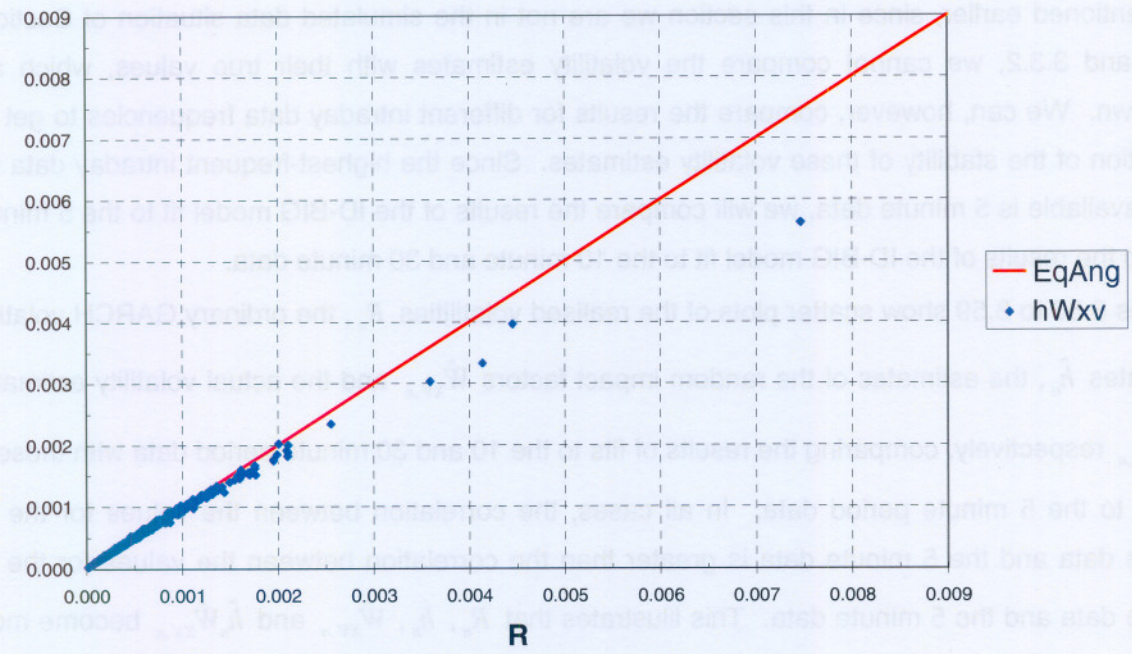

Figure 3.53: Scatter plot comparing $\hat{h}_n \hat{W}_{XV,n}$ and R_n for ID-BIG5 fit to IBM data.

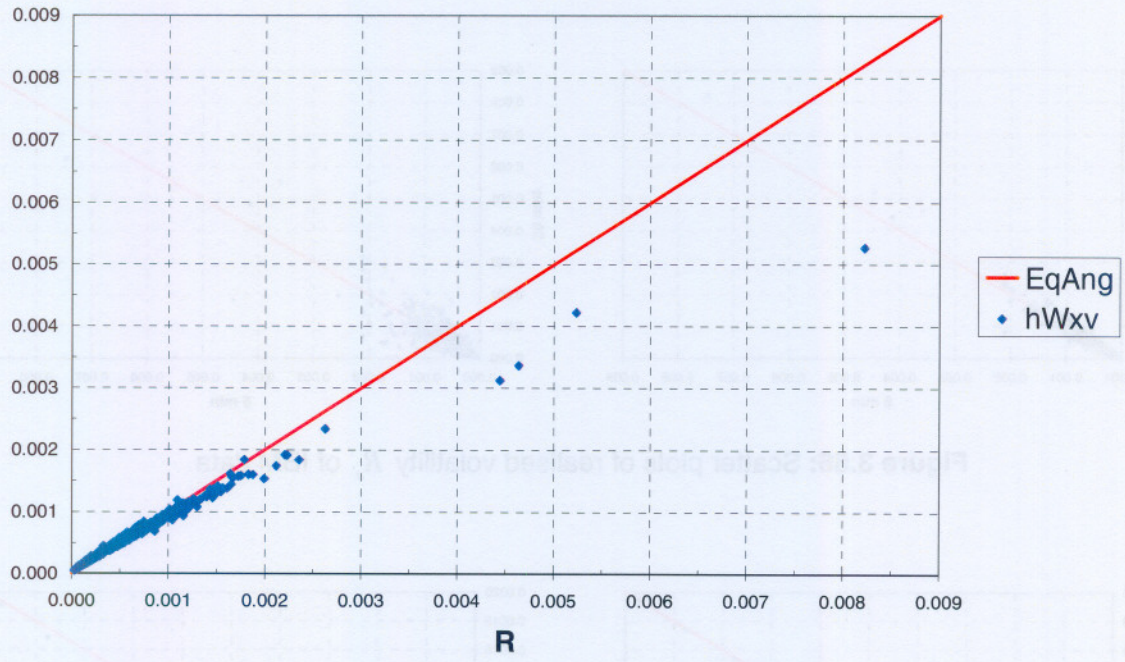

Figure 3.54: Scatter plot comparing $\hat{h}_n \hat{W}_{XV,n}$ and R_n for ID-BIG10 fit to IBM data.

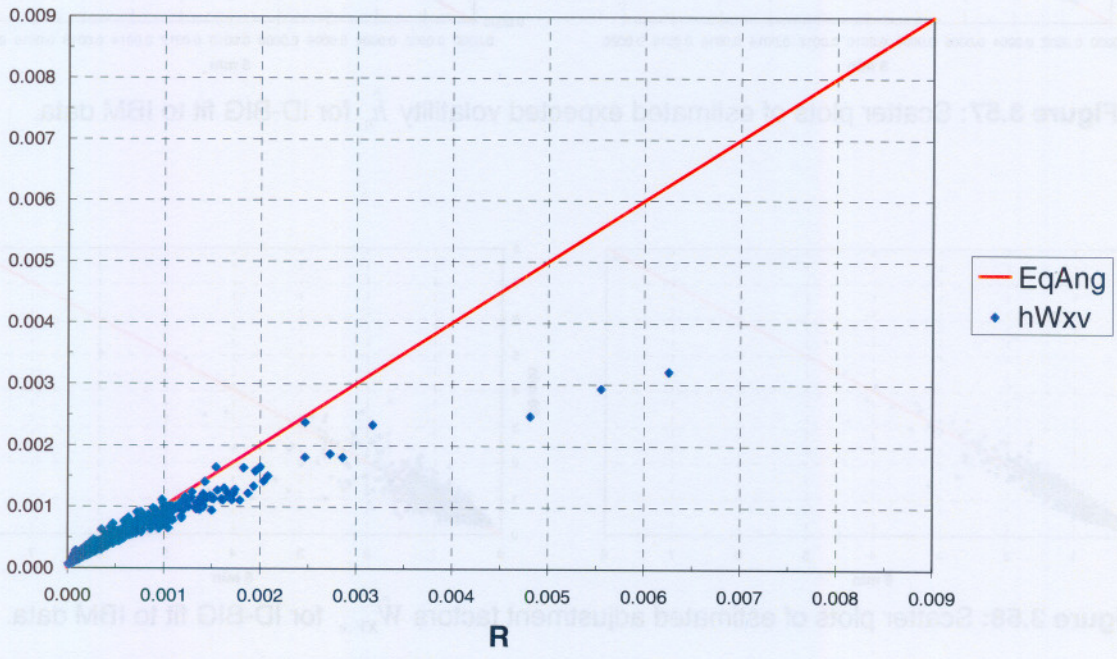

Figure 3.55: Scatter plot comparing $\hat{h}_n \hat{W}_{XV,n}$ and R_n for ID-BIG30 fit to IBM data.

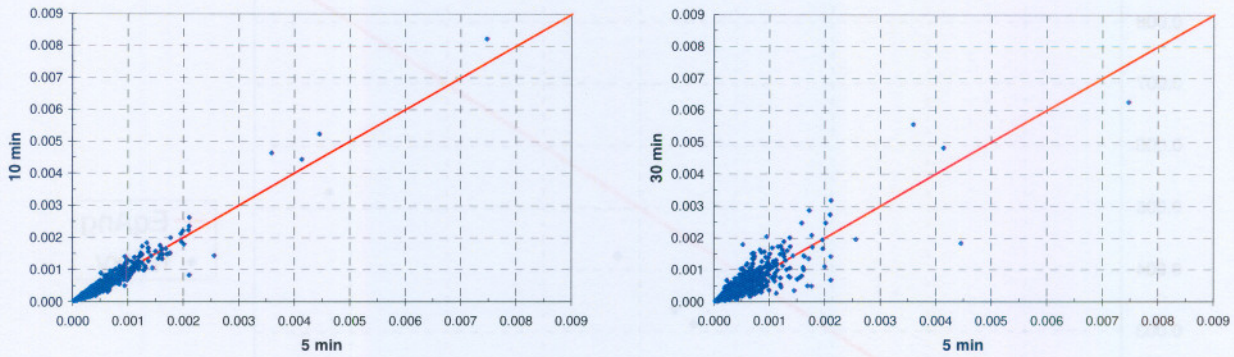

Figure 3.56: Scatter plots of realised volatility R_n of IBM data.

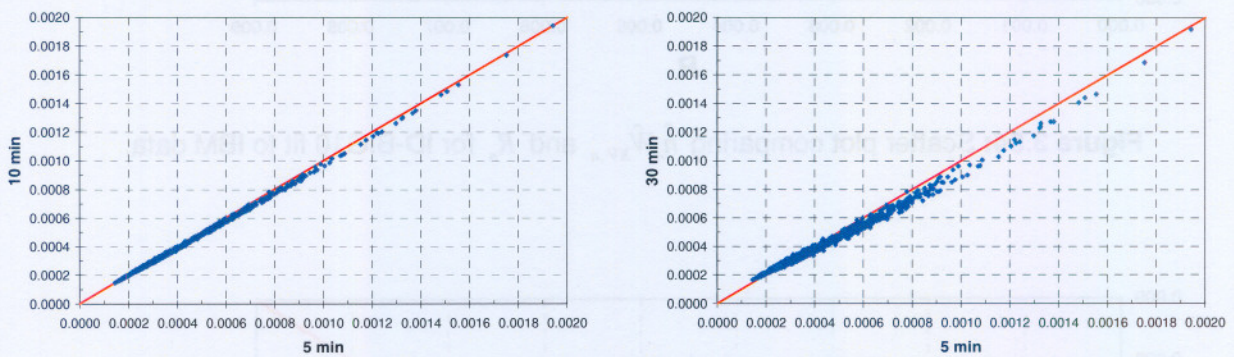

Figure 3.57: Scatter plots of estimated expected volatility $\hat{h}_n$ for ID-BIG fit to IBM data.

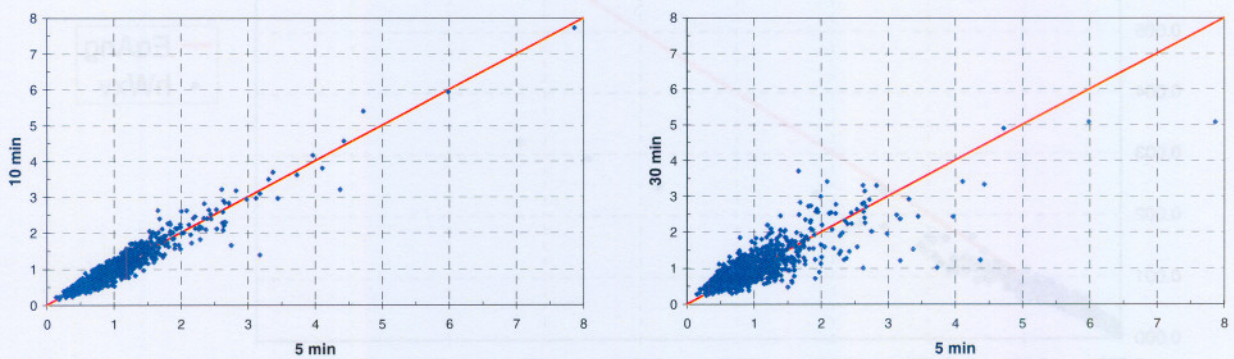

Figure 3.58: Scatter plots of estimated adjustment factors $\hat{W}_{XV,n}$ for ID-BIG fit to IBM data.

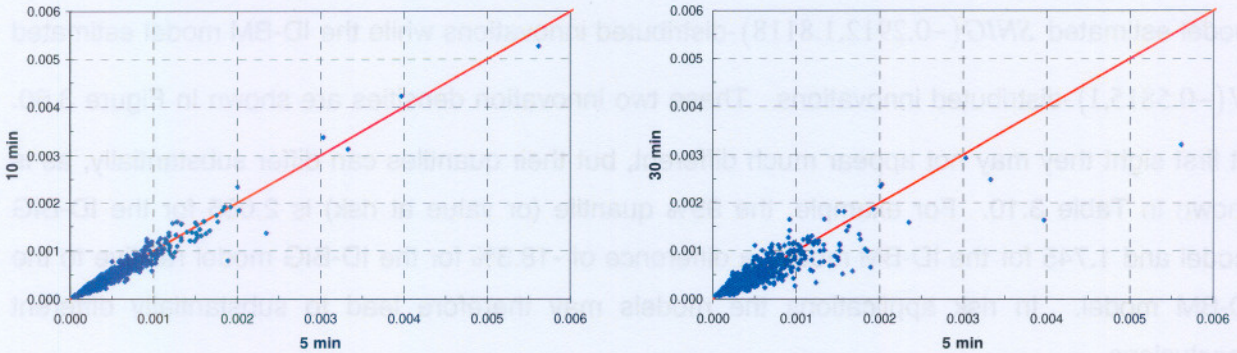

Figure 3.59: Scatter plots of estimated actual volatility $\hat{h}_n \hat{W}_{XV,n}$ for ID-BIG fit to IBM data.

Regarding $\hat{h}_n$, we see from Figure 3.57 that the 5 and 10 minute values are virtually identical while the agreement between the 5 and 30 minute values is still quite close, although not as close as for the 5 and 10 minute values. For R_n, $\hat{W}_{XV,n}$ and $\hat{h}_n \hat{W}_{XV,n}$ the agreement between the 5, 10 and 30 minute data (as seen in Figures 3.56, 3.58, and 3.59) is fairly close, although not as close as for $\hat{h}_n$. This illustrates that there is not a sizeable change in the accuracy with which the ordinary GARCH volatility estimator $\hat{h}_n$ estimates the unknown underlying actual volatility when using more and more intraday data. However, the changes in realised volatility R_n for different intraday data frequencies are more prominent and, as we have already pointed out in Section 2.5.3, the accuracy with which the adjusted volatility estimator $\hat{h}_n \hat{W}_{XV,n}$ estimates the actual volatility improves with an increase in the number of periods, i.e. with shorter period lengths.

These results illustrate that the major benefit of having higher frequency data is the greater accuracy with which one is able to estimate actual volatility. Furthermore, the fact that the volatility estimates obtained from the ID-BIG fits to 5, 10 and 30 minute data do not differ drastically supports the hunch (seen in Table 3.8) that the parameter estimates results should be quite similar if the data of different period lengths are used.

Before moving on to the DW, Q and LM tests, we will now briefly examine the innovation distributions estimated by the ID-BIG and ID-BM model fits to the 5 minute IBM data. The ID-BIG model estimated $SNIG(-0.2912, 1.8118)$-distributed innovations while the ID-BM model estimated $N(-0.5815, 1)$-distributed innovations. These two innovation densities are shown in Figure 3.60. At first sight they may not appear much different, but their quantiles can differ substantially, as is shown in Table 3.10. For example, the 99% quantile (or value at risk) is 2.065 for the ID-BIG model and 1.745 for the ID-BM model, a difference of -18.3% for the ID-BIG model relative to the ID-BM model. In risk applications the models may therefore lead to substantially different conclusions.

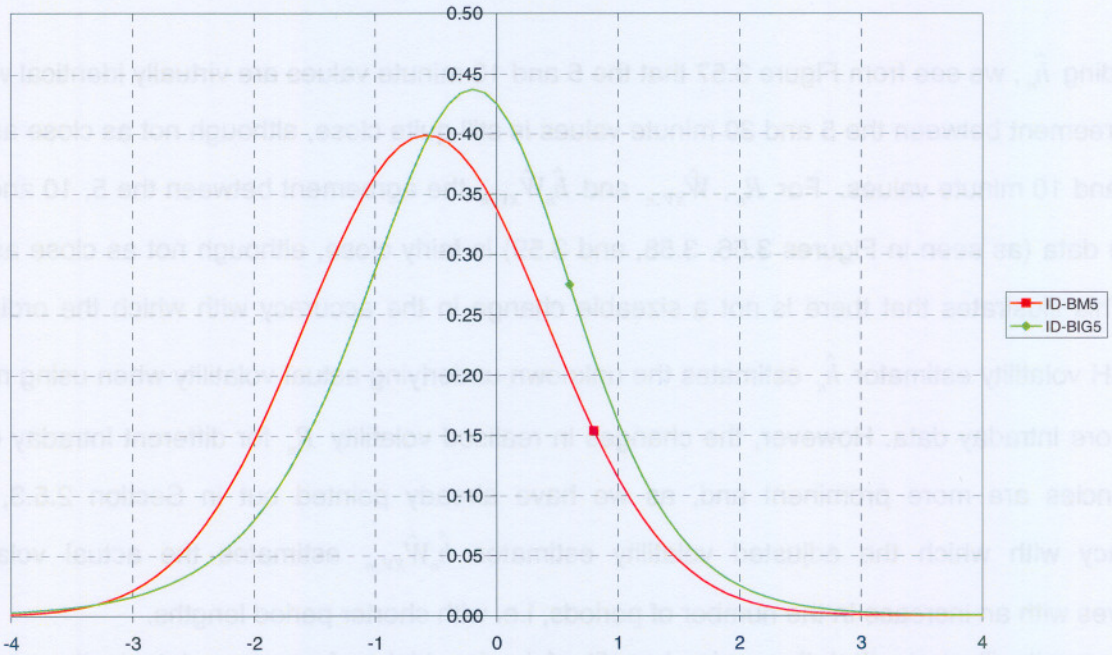

Figure 3.60: Innovation distributions of ID-BM5 and ID-BIG5 fits to IBM data.

Quantile	ID-BM5	ID-BIG5	Difference
1%	-2.908	-2.977	-2.4%
5%	-2.226	-2.000	10.2%
10%	-1.863	-1.554	16.6%
25%	-1.256	-0.898	28.5%
50%	-0.581	-0.255	56.1%
75%	0.093	0.356	-282.8%
90%	0.700	0.926	-32.3%
95%	1.063	1.294	-21.7%
99%	1.745	2.065	-18.3%

Table 3.10: Innovation distribution quantiles of ID-BM5 and ID-BIG5 fits to IBM data.

Testing for dependence

Finally, consider the results of the DW, Q and LM tests on the residuals and estimates obtained from fitting the various models to the IBM data (as argued before we perform the tests on the natural logarithm of $\hat{W}_{X,n}$, $\hat{W}_{XAB,n}$, $\hat{W}_{XV,n}$, $\hat{V}_n$ and $\hat{C}_{XV,n}$), given in Tables 3.11 and 3.12 below.

As seen Table 3.11, for the D-NOR, D-NIG, HLC-BM and HLC-BIG model fits to the IBM data the tests indicate that the residuals and estimates conform to the requirements of independence with no remaining ARCH effects in all the cases except for $\log \hat{W}_{XAB,n}$. The DW test strongly indicates the presence of positive autocorrelation in the $\log \hat{W}_{XAB,n}$ estimates. Going on to really high-frequency data we see from Table 3.12 that that there are more exceptions. For the ID-BM and ID-BIG models, the tests indicate that the residuals and estimates conform to the requirements of independence with no remaining ARCH effects in all the cases except the following. The DW test strongly indicates positive autocorrelation for $\log \hat{V}_n$, $\log \hat{W}_{XV,n}$ and $\log \hat{C}_{XV,n}$ (remember that $V_n = W_n C_n$) and the Q and LM tests indicate remaining (mainly second order) ARCH effects for $\log \hat{V}_n$ and $\log \hat{W}_{XV,n}$. The difference in results for $\log \hat{W}_{X,n}$, $\log \hat{W}_{XAB,n}$ and $\log \hat{W}_{XV,n}$ can be attributed to the differing degree of accuracy with which they estimate $\log W_n$. We have seen that $\log \hat{W}_{XV,n}$ is the most accurate estimate of the three (since it makes use of full intraday data) and we therefore focus on the dependence indicated for $\log \hat{W}_{XV,n}$.

In order to examine the nature of this dependence we fit an $AR(1)$ (the simplest) model to the $\log \hat{W}_{XV,n}$ estimates obtained from the ID-BIG model fits to the 5, 10 and 30 minute data and then test for remaining dependence. The results appear in Table 3.13. Firstly, we see that there is now no indication of remaining ARCH effects. Secondly, the first order autocorrelation is estimated at 0.391, 0.389 and 0.242 for the 5, 10 and 30 minute data respectively and there is no indication of remaining autocorrelation. Thus, it seems as if an $AR(1)$ process is sufficient to model the dependence in the volatility adjustment factors W_n.

Effects of dependence

Two natural questions to ask at this stage are "how can we explain this dependence?" and "what effect will it have on our analyses?" Firstly, we offer a common sense practical explanation.

When we think about the practical meaning of the volatility adjustment factors W_n, it becomes clear that dependence in these factors is to be expected. In Section 3.2.2 we argued that W_n represents the impact of news and events occurring after close of market on day $n-1$ and effects of trading on other markets around the world. The ordinary GARCH volatility estimate $\hat{h}_n$ is unable to incorporate these impacts, which is why we adjust $\hat{h}_n$ to a new volatility estimator ($\hat{h}_n\hat{W}_{XV,n}$ in the ID-BIG case) that estimates the actual volatility much more accurately. However, news and events that happen shortly after close of market on day $n-1$ will have a more profound impact on trading during day n than will news and events that happen shortly before close of market on day n, which will actually have more of an impact on trading during day $n+1$, after it has been properly digested by the players in the market. News and events that occur close to the middle of the trading day may have a considerable effect on trading during both days n and $n+1$. Thus, there will be a trade-off between the amount of time elapsed after close of market on day $n-1$ until the event takes place and the intensity of the impact it may have on trading on day n as opposed to day $n+1$. As W_n represents the total impact of all such impact factors occurring from close of market on day $n-1$ up until close of market on day n it is only natural for the series $\{W_n\}$ to be time dependent.

We can also expect this dependence to have an impact on the results obtained from fitting the models assuming no dependence for data in which the dependence is actually present. This causes a dilemma since, for all we know at this stage, the results may be so affected as to be quite inaccurate and misleading. The only way to overcome this dilemma is to extend the ID-BIG model to accommodate this dependence, to fit the extended model to our data and to compare the results with those obtained under the independence assumption.

D-NOR	X=Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.054	0.817	0.057	0.811
2	0.250	0.883	0.228	0.892
3	1.729	0.631	1.840	0.606
4	1.731	0.785	1.855	0.762
5	2.392	0.793	2.490	0.778
6	2.464	0.872	2.606	0.856
7	4.382	0.735	4.349	0.739
8	4.505	0.809	4.519	0.808
9	5.125	0.823	5.282	0.809
10	5.142	0.882	5.290	0.871
Durbin-Watson				2.032
Pr<DW				0.692
Pr>DW				0.308

D-NIG	Z			
Order	Q	Pr>Q	LM	Pr>LM
1	0.057	0.811	0.064	0.801
2	0.316	0.854	0.273	0.873
3	2.272	0.518	2.507	0.474
4	2.272	0.686	2.518	0.642
5	2.805	0.730	2.992	0.701
6	2.831	0.830	3.071	0.800
7	2.947	0.890	3.137	0.872
8	3.003	0.934	3.208	0.921
9	3.433	0.945	3.766	0.926
10	3.433	0.969	3.768	0.957
Durbin-Watson				2.010
Pr<DW				0.561
Pr>DW				0.439

D-NIG	X			
Order	Q	Pr>Q	LM	Pr>LM
1	0.143	0.706	0.148	0.701
2	0.201	0.905	0.192	0.909
3	1.956	0.582	2.087	0.555
4	1.986	0.738	2.151	0.708
5	2.467	0.781	2.591	0.763
6	2.554	0.862	2.725	0.843
7	4.369	0.737	4.375	0.736
8	4.476	0.812	4.531	0.806
9	5.135	0.822	5.290	0.808
10	5.145	0.881	5.293	0.871
Durbin-Watson				2.013
Pr<DW				0.580
Pr>DW				0.420

D-NIG	log Wx			
Order	Q	Pr>Q	LM	Pr>LM
1	0.423	0.515	0.421	0.517
2	0.432	0.806	0.432	0.806
3	0.932	0.818	0.945	0.815
4	0.995	0.911	1.024	0.906
5	1.030	0.960	1.054	0.958
6	1.320	0.971	1.351	0.969
7	11.900	0.104	11.706	0.111
8	11.909	0.155	11.708	0.165
9	12.180	0.203	11.970	0.215
10	13.024	0.222	12.597	0.247
Durbin-Watson				2.005
Pr<DW				0.534
Pr>DW				0.466

HLC-BM	X=Z			
Order	Q	Pr>Q	LM	Pr>LM
1	1.348	0.246	1.357	0.244
2	1.359	0.507	1.385	0.500
3	4.255	0.235	4.485	0.214
4	4.538	0.338	4.976	0.290
5	4.649	0.460	5.031	0.412
6	4.764	0.575	5.226	0.515
7	5.754	0.569	6.010	0.539
8	5.784	0.672	6.072	0.639
9	6.533	0.686	6.862	0.652
10	6.536	0.768	6.867	0.738
Durbin-Watson				2.036
Pr<DW				0.715
Pr>DW				0.285

HLC-BIG	Z			
Order	Q	Pr>Q	LM	Pr>LM
1	1.158	0.282	1.178	0.278
2	1.164	0.559	1.202	0.548
3	4.520	0.211	4.968	0.174
4	4.742	0.315	5.441	0.245
5	4.755	0.447	5.442	0.364
6	4.813	0.568	5.614	0.468
7	4.815	0.683	5.624	0.584
8	4.815	0.777	5.625	0.689
9	5.106	0.825	6.032	0.737
10	5.112	0.884	6.033	0.813
Durbin-Watson				2.001
Pr<DW				0.504
Pr>DW				0.496

HLC-BIG	X			
Order	Q	Pr>Q	LM	Pr>LM
1	1.573	0.210	1.580	0.209
2	1.624	0.444	1.666	0.435
3	4.855	0.183	5.147	0.161
4	5.205	0.267	5.753	0.218
5	5.254	0.386	5.763	0.330
6	5.444	0.488	6.079	0.414
7	6.212	0.515	6.637	0.468
8	6.223	0.622	6.662	0.574
9	7.055	0.631	7.563	0.579
10	7.056	0.720	7.564	0.671
Durbin-Watson				1.999
Pr<DW				0.495
Pr>DW				0.505

HLC-BIG	log Wx			
Order	Q	Pr>Q	LM	Pr>LM
1	1.740	0.187	1.731	0.188
2	1.843	0.398	1.871	0.393
3	3.220	0.359	3.321	0.345
4	3.339	0.503	3.519	0.475
5	3.346	0.647	3.537	0.618
6	3.609	0.730	3.855	0.696
7	9.732	0.204	9.667	0.208
8	9.871	0.274	9.698	0.287
9	10.281	0.328	10.088	0.343
10	11.143	0.347	10.712	0.380
Durbin-Watson				2.061
Pr<DW				0.833
Pr>DW				0.167

HLC-BIG	log Wxab			
Order	Q	Pr>Q	LM	Pr>LM
1	0.121	0.728	0.123	0.726
2	3.763	0.152	3.780	0.151
3	4.013	0.260	4.055	0.256
4	4.434	0.350	4.589	0.332
5	4.753	0.447	4.792	0.442
6	5.453	0.487	5.280	0.508
7	5.474	0.602	5.281	0.626
8	6.509	0.590	6.565	0.584
9	7.552	0.580	7.621	0.573
10	8.083	0.621	8.335	0.596
Durbin-Watson				1.594
Pr<DW				0.000
Pr>DW				1.000

Table 3.11: DW, Q and LM tests for daily and HLC model fits to IBM data.

ID-BM5	X=Z				ID-BM5	log V				ID-BM5	T			
Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM
1	1.611	0.204	1.607	0.205	1	3.441	0.064	3.467	0.063	1	0.876	0.349	0.877	0.349
2	1.743	0.418	1.790	0.409	2	6.013	0.050	5.749	0.057	2	0.890	0.641	0.879	0.644
3	3.978	0.264	4.262	0.235	3	6.052	0.109	5.750	0.124	3	1.249	0.741	1.388	0.708
4	4.572	0.334	5.134	0.274	4	7.021	0.135	6.599	0.159	4	1.268	0.867	1.467	0.832
5	4.620	0.464	5.141	0.399	5	7.063	0.216	6.685	0.245	5	2.594	0.762	2.559	0.768
6	4.631	0.592	5.184	0.521	6	7.078	0.314	6.710	0.349	6	2.920	0.819	2.851	0.827
7	5.106	0.647	5.533	0.595	7	7.362	0.392	7.074	0.421	7	2.971	0.888	2.880	0.896
8	5.189	0.737	5.644	0.687	8	10.024	0.263	9.460	0.305	8	3.109	0.927	2.989	0.935
9	5.834	0.756	6.241	0.716	9	10.300	0.327	9.549	0.388	9	3.186	0.957	3.027	0.963
10	5.849	0.828	6.256	0.793	10	10.709	0.381	10.240	0.420	10	4.177	0.939	4.319	0.932
Durbin-Watson				2.028	Durbin-Watson				1.156	Durbin-Watson				1.935
Pr<DW				0.668	Pr<DW				0.000	Pr<DW				0.152
Pr>DW				0.332	Pr>DW				1.000	Pr>DW				0.849
ID-BIG5	X				ID-BIG5	log V				ID-BIG5	T			
Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM
1	2.192	0.139	2.194	0.139	1	3.060	0.080	3.083	0.079	1	0.654	0.419	0.662	0.416
2	2.333	0.311	2.407	0.300	2	6.314	0.043	6.053	0.049	2	0.677	0.713	0.669	0.716
3	6.117	0.106	6.525	0.089	3	6.469	0.091	6.103	0.107	3	1.300	0.729	1.509	0.680
4	6.586	0.160	7.360	0.118	4	7.923	0.095	7.321	0.120	4	1.436	0.838	1.789	0.775
5	6.591	0.253	7.368	0.195	5	8.059	0.153	7.566	0.182	5	2.354	0.798	2.478	0.780
6	6.860	0.334	7.838	0.250	6	8.124	0.229	7.659	0.264	6	2.778	0.836	2.815	0.832
7	7.249	0.403	8.048	0.328	7	8.315	0.306	7.941	0.338	7	2.820	0.901	2.829	0.900
8	7.250	0.510	8.049	0.429	8	10.050	0.262	9.524	0.300	8	3.034	0.932	2.995	0.935
9	8.191	0.515	9.114	0.427	9	10.301	0.327	9.646	0.380	9	3.065	0.962	3.006	0.964
10	8.204	0.609	9.140	0.519	10	10.961	0.361	10.605	0.389	10	3.844	0.954	4.040	0.946
Durbin-Watson				1.987	Durbin-Watson				1.156	Durbin-Watson				1.938
Pr<DW				0.416	Pr<DW				0.000	Pr<DW				0.162
Pr>DW				0.584	Pr>DW				1.000	Pr>DW				0.838
ID-BIG5	log Wxv				ID-BIG5	Zxv				ID-BIG5	log Cxv			
Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM	Order	Q	Pr>Q	LM	Pr>LM
1	3.508	0.061	3.534	0.060	1	1.524	0.217	1.536	0.215	1	0.073	0.787	0.075	0.785
2	7.318	0.026	6.990	0.030	2	1.566	0.457	1.624	0.444	2	0.212	0.899	0.217	0.897
3	7.506	0.057	7.046	0.071	3	2.451	0.484	2.804	0.423	3	0.296	0.961	0.300	0.960
4	9.505	0.050	8.720	0.069	4	2.856	0.582	3.492	0.479	4	0.390	0.983	0.393	0.983
5	9.648	0.086	9.009	0.109	5	3.848	0.572	4.156	0.527	5	0.489	0.993	0.493	0.992
6	9.664	0.140	9.050	0.171	6	4.032	0.672	4.287	0.638	6	1.001	0.986	0.989	0.986
7	9.983	0.190	9.474	0.220	7	4.043	0.775	4.288	0.746	7	1.197	0.991	1.156	0.992
8	11.761	0.162	10.996	0.202	8	4.291	0.830	4.497	0.810	8	1.764	0.987	1.747	0.988
9	12.086	0.209	11.155	0.265	9	4.397	0.883	4.581	0.869	9	1.781	0.995	1.765	0.995
10	12.609	0.246	11.995	0.285	10	5.010	0.891	5.367	0.865	10	2.089	0.996	2.081	0.996
Durbin-Watson				1.156	Durbin-Watson				1.931	Durbin-Watson				1.276
Pr<DW				0.000	Pr<DW				0.136	Pr<DW				0.000
Pr>DW				1.000	Pr>DW				0.864	Pr>DW				1.000

Table 3.12: DW, Q and LM tests for intraday model fits to IBM data.

IBM	ID-BIG5				ID-BIG10				ID-BIG30			
	Durbin-Watson			1.156	Durbin-Watson			1.156	Durbin-Watson			1.156
	Pr<DW			0.000	Pr<DW			0.000	Pr<DW			0.000
	Pr>DW			1.000	Pr>DW			1.000	Pr>DW			1.000
	AR(1) Autocorrelation			0.391	AR(1) Autocorrelation			0.389	AR(1) Autocorrelation			0.242
	Standard error			0.028	Standard error			0.029	Standard error			0.031
Order	**Q**	**Pr>Q**	**LM**	**Pr>LM**	**Q**	**Pr>Q**	**LM**	**Pr>LM**	**Q**	**Pr>Q**	**LM**	**Pr>LM**
1	0.604	0.437	0.591	0.442	0.066	0.798	0.063	0.802	2.651	0.104	2.602	0.107
2	4.425	0.109	4.378	0.112	0.513	0.774	0.526	0.769	4.241	0.120	4.040	0.133
3	4.433	0.218	4.379	0.223	0.527	0.913	0.550	0.908	4.541	0.209	4.191	0.242
4	4.911	0.297	4.728	0.316	2.298	0.681	2.318	0.678	5.411	0.248	5.162	0.271
5	5.370	0.372	5.103	0.403	3.173	0.673	3.116	0.682	7.071	0.215	6.836	0.233
6	5.555	0.475	5.353	0.499	4.384	0.625	4.413	0.621	7.318	0.293	7.097	0.312
7	5.680	0.578	5.529	0.596	4.385	0.735	4.417	0.731	7.319	0.396	7.108	0.418
8	6.252	0.619	6.219	0.623	4.587	0.801	4.656	0.794	8.497	0.387	8.363	0.399
9	7.279	0.608	7.319	0.604	6.632	0.675	7.072	0.630	10.983	0.277	11.117	0.268
10	7.282	0.699	7.319	0.695	6.635	0.759	7.096	0.716	11.290	0.335	11.344	0.331
Durbin-Watson				2.000	Durbin-Watson			2.097	Durbin-Watson			2.056
Pr<DW				0.495	Pr<DW			0.936	Pr<DW			0.809
Pr>DW				0.505	Pr>DW			0.064	Pr>DW			0.191

Table 3.13: DW, Q and LM tests for $\log \hat{W}_{XV,n}$ of ID-BIG5/10/30 fits to IBM data.

3.4 Conclusion

In this chapter we extended the basic daily normal GARCH model by assuming NIG-distributed innovations (or BIG processes) instead of normally distributed innovations (or BM processes) and by catering for the use of high, low and close prices, as well as full intraday data. We found that the models allowing for NIG-distributed innovations (with heavier tails) and catering for full intraday data fit empirical data sets better than models assuming normally distributed innovations and making use of only daily data.

However, we found indications of dependence in the random impact factors, which contradicts our assumption of independence. The influence of this dependence on the reliability of the results obtained through fitting these models will be considered in Chapter 4, where we further extend the ID-BIG model to cater for this dependence.

CHAPTER 4

GARCH MODELS FOR TEMPORALLY DEPENDENT BROWNIAN INVERSE GAUSSIAN PROCESSES

4.1 Introduction

When we examined the results of fitting the ID-BIG model to the IBM data, towards the end of Section 3.3.3, we encountered empirical evidence of time dependence in the random impact factors W_n, which we will focus on next. In Section 4.2 we will extend the NIG innovation based models and introduce models that allow for such time dependence. Since the time dependence was observed when fitting the ID-BIG model we start by extending the ID-BIG model to the TD-ID-BIG model. Extensions to the D-NIG and HLC-BIG models will receive attention towards the end of Section 4.2, where we will argue that the absence of intraday data causes the estimates of W_n, and thus the estimate of the dependence in W_n, to be extremely inaccurate, rendering these extensions fruitless. In Section 4.3 we fit the TD-ID-BIG model to both generated and empirical data sets, perform residual analysis and report the findings. We find that the time-dependent model fits empirical data well, as long as enough intraday data is available.

4.2 Extending the ID-BIG Model

We will use the same notation and assumptions as for the ID-BIG model (see (3.2.12)) introduced in Section 3.2.3 with one exception, namely that we no longer assume that the W_n's are independent over n. We still want the W_n's to be $UIG(\psi)$-distributed marginally, but over time n they must form at least an $AR(1)$ process. Unfortunately no suitable multivariate GIG distribution theory seems to exist that can be used to specify a model for the W_n's that meets these requirements. We therefore proceed via a normally distributed $AR(1)$ process and appropriate transformations as follows. Let $\tilde{U}_1,\ldots,\tilde{U}_N$ be iid $N(0,1)$ and define

$$U_1 = \tilde{U}_1 \text{ and } U_n = \rho U_{n-1} + \sqrt{1-\rho^2}\tilde{U}_n \text{ , } n = 2,\ldots,N \text{ .} \qquad (4.2.1)$$

$\{U_n\}$ now follows an $AR(1)-N(0,1)$ process, with

$$\begin{aligned} E(U_n) &= 0 \\ Var(U_n) &= \rho^2 Var(U_{n-1}) + (1-\rho^2) Var(\tilde{U}_n) = \rho^2 + 1 - \rho^2 = 1 \\ Cov(U_n, U_{n-1}) &= \rho \ . \end{aligned} \tag{4.2.2}$$

Let $F_{UIG(\psi)}(w)$ be the $UIG(\psi)$ distribution function and $F^{-1}_{UIG(\psi)}(p)$ the inverse of the $UIG(\psi)$ distribution function. Define

$$W_n = F^{-1}_{UIG(\psi)}(\Phi(U_n)) = g(U_n) \tag{4.2.3}$$

using $g(U_n)$ to simplify notation. Then we call the joint distribution of $W_1,\ldots,W_N$ the $AR(\rho)-UIG(\psi)$ distribution. Notice that, since $U_n \sim N(0,1)$, $\Phi(U_n) \sim U(0,1)$ and therefore $W_n \sim UIG(\psi)$. So, each marginal of $(W_1,\ldots,W_N)$ is $UIG(\psi)$-distributed as required. However, since the U_n's have lag1 autocorrelation the W_n's also have lag1 autocorrelation, say $\rho_w = Corr(W_n, W_{n-1}) = \rho_w(\rho,\psi)$. In Chapter 3 we motivated performing the Durbin Watson (DW) test for autocorrelation on the $\log \hat{W}_n$'s instead of the $\hat{W}_n$'s and we shall do the same when examining the results of fitting this time dependent version of the ID-BIG model in Section 4.3. Of course, taking logarithms, the $\log W_n$'s also have lag1 autocorrelation, say $\rho_{\log w} = Corr(\log W_n, \log W_{n-1}) = \rho_{\log w}(\rho,\psi)$. Explicit expressions for these lag1 autocorrelation functions are unavailable but we approximate them in Section A.2 of the Appendix, where we argue that the autocorrelation in the $\log W_n$'s is approximately equal to the autocorrelation in the original $N(0,1)$-distributed $\{U_n\}$, i.e. $\rho_{\log w} \approx \rho$.

We will indicate this ID-BIG model with the $AR(\rho)-UIG(\psi)$ distributional assumption by TD-ID-BIG, where TD refers to the time dependence in the W_n's, ID again refers to intraday data and BIG again refers to the Brownian inverse Gaussian process; thus the "time dependent intraday Brownian inverse Gaussian" model. In this case it is much more difficult to calculate the likelihood function. To illustrate the problem, let us ignore the intraday data for a while and look at what happens when allowing for this lag1 autocorrelation in W_n in the simpler case of the D-NIG model of Section 3.2.2. From (3.2.7) we then have $Y_n = \mu_n + \sqrt{h_n}\beta W_n + \sqrt{h_n W_n} Z_n$, so that

$$\frac{Y_n - \mu_n}{\sqrt{h_n}} = \beta W_n + \sqrt{W_n} Z_n \ . \tag{4.2.4}$$

In the models up to Chapter 3 we assumed that W_n and Z_n were independent of $Y_1,\ldots,Y_{n-1}$, so that the conditional distribution of $\beta W_n + \sqrt{W_n} Z_n$ given $Y_1,\ldots,Y_{n-1}$ is the same as the unconditional distribution of $\beta W_n + \sqrt{W_n} Z_n$, which is $NIG(\beta,\psi)$. Hence, $(Y_n - \mu_n)/\sqrt{h_n} \,|Y_1,\ldots,Y_{n-1} \sim NIG(\beta,\psi)$ and therefore

$$Y_n | Y_1,\ldots Y_{n-1} \sim \frac{1}{\sqrt{h_n}} f_{NIG(\beta,\psi)}\left(\frac{y_n - \mu_n}{\sqrt{h_n}}\right) . \tag{4.2.5}$$

By chain rule decomposition, we could then write

$$\begin{aligned} f_{Y_1,\ldots,Y_N}(y_1,\ldots,y_N) &= f_{Y_1}(y_1) f_{Y_2|Y_1}(y_2|y_1)\cdots f_{Y_N|Y_{N-1}}(y_N|y_{N-1}) \\ &= \prod_{n=1}^{N} \frac{1}{\sqrt{h_n}} f_{NIG(\beta,\psi)}\left(\frac{y_n - \mu_n}{\sqrt{h_n}}\right) \end{aligned} \tag{4.2.6}$$

which gives us the joint density function for $Y_1,\ldots,Y_N$ (and therefore the likelihood function) when there is no dependence in the W_n's. Likelihood based statistical analysis then becomes possible. Having such an explicit expression for the likelihood function of traditional GARCH models is one of the great advantages of dealing with GARCH models as opposed to stochastic volatility models, for which this is usually not the case (see Danielsson (1994)).

When we allow for lag1 autocorrelation in $\{W_n\}$, W_n depends on W_{n-1} and, since $Y_{n-1} = \mu_{n-1} + \sqrt{h_{n-1}}\beta W_{n-1} + \sqrt{h_{n-1} W_{n-1}} Z_{n-1}$ we see that Y_{n-1} also depends on W_{n-1}. This means that we cannot conclude that W_n is independent of Y_{n-1} and therefore it is no longer true that the conditional distribution of $\beta W_n + \sqrt{W_n} Z_n$ given $Y_1,\ldots,Y_{n-1}$ is the same as the unconditional distribution of $\beta W_n + \sqrt{W_n} Z_n$, i.e. we cannot conclude that $(Y_n - \mu_n)/\sqrt{h_n} \,|Y_1,\ldots,Y_{n-1} \sim NIG(\beta,\psi)$. The easy availability of the likelihood function is lost and statistical inference based on the likelihood function is more difficult. This is a difficulty that stochastic volatility models also have to deal with.

Some of the estimation techniques that have been proposed to overcome this difficulty are the Generalised Method of Moments used by Melino and Turnbull (1990) and Markov Chain Monte Carlo procedures used by Jacquier, Polson and Rossi (1999). See Ghysels, Harvey and Renault (1996) for a survey of these estimation techniques and Andersen, Chung and Sørensen (1999) for performance comparisons. Since we have worked with likelihood based inference, our preference

is to continue along this line. The question then becomes one of how we can compute or approximate the likelihood function in this more general context. We follow a method proposed by Liesenfeld and Richard (2003) to obtain the likelihood function by using the Efficient Importance Sampling (EIS) method as follows.

Efficient Importance Sampling

It is still possible to use the chain rule decomposition if we include the latent factor U_n and then derive the joint density of the observed data and the latent factors. Setting, as in (2.7.20) and Section 3.2.3, $V_n = \tilde{V}_n / h_n = \left(IR_n - Y_n^2\right)/h_n = W_n C_n$, it follows from our assumptions and (1.6.4) that Y_n given $(Y_1,V_1,U_1),\ldots,(Y_{n-1},V_{n-1},U_{n-1}),U_n$ is $N\left(\mu_n + \beta\sqrt{h_n}\,g(U_n), h_n g(U_n)\right)$-distributed, that $V_n/g(U_n)$ given $(Y_1,V_1,U_1),\ldots,(Y_{n-1},V_{n-1},U_{n-1}),U_n$ is χ_ι^2-distributed independently of Y_n and that U_n given $(Y_1,V_1,U_1),\ldots,(Y_{n-1},V_{n-1},U_{n-1})$ is $N\left(\rho U_{n-1}, 1-\rho^2\right)$-distributed, with $\iota = I-1$ and I the number of intervals per day. The joint conditional density function of the observed data and the latent factors (Y_n,V_n,U_n) in the arguments (y_n,v_n,u_n), given $(Y_1,V_1,U_1),\ldots,(Y_{n-1},V_{n-1},U_{n-1}) = (y_1,v_1,u_1),\ldots,(y_{n-1},v_{n-1},u_{n-1})$, can therefore be written as

$$
\begin{aligned}
& f\left(y_n,v_n,u_n \middle| (y_1,v_1,u_1),\ldots,(y_{n-1},v_{n-1},u_{n-1})\right) \\
&= f\left(y_n,v_n \middle| (y_1,v_1,u_1),\ldots,(y_{n-1},v_{n-1},u_{n-1}),u_n\right) f\left(u_n \middle| (y_1,v_1,u_1),\ldots,(y_{n-1},v_{n-1},u_{n-1})\right) \\
&= f\left(y_n \middle| (y_1,v_1,u_1),\ldots,(y_{n-1},v_{n-1},u_{n-1}),u_n\right) f\left(v_n \middle| (y_1,v_1,u_1),\ldots,(y_{n-1},v_{n-1},u_{n-1}),u_n\right) \\
&\qquad \cdot f\left(u_n \middle| (y_1,v_1,u_1),\ldots,(y_{n-1},v_{n-1},u_{n-1})\right) \\
&= f_n(y_n,v_n,u_n)\, p_n\left(u_n \middle| u_{n-1}\right)
\end{aligned}
\tag{4.2.7}
$$

where we set, for ease of notation,

$$
\begin{aligned}
f_n(y,v,u) &= \frac{1}{\sqrt{h_n g(u)}}\varphi\left(\frac{y-\mu_n-\beta\sqrt{h_n}\,g(u)}{\sqrt{h_n g(u)}}\right)\frac{1}{g(u)} f_{\chi_\iota^2}\left(\frac{v}{g(u)}\right) \\
p_n\left(u \middle| s\right) &= \frac{1}{\sqrt{1-\rho_n^2}}\varphi\left(\frac{u-\rho_n s}{\sqrt{1-\rho_n^2}}\right),\quad \rho_n = \begin{cases}\rho & \text{if } n>1 \\ 0 & \text{if } n=1\end{cases}.
\end{aligned}
\tag{4.2.8}
$$

Note that $(y_1,v_1,u_1),\ldots,(y_{n-1},v_{n-1},u_{n-1})$ are subsumed in μ_n and h_n via the definitions of μ_n and h_n.

The joint density function of $(Y_1,V_1,U_1),\ldots,(Y_N,V_N,U_N)$ in the arguments $(y_1,v_1,u_1),\ldots,(y_N,v_N,u_N)$ is then given by

$$f\left((y_1,v_1,u_1),\ldots,(y_N,v_N,u_N)\right)=\prod_{n=1}^{N} f_n\left(y_n,v_n,u_n\right)p_n\left(u_n|u_{n-1}\right)\ . \tag{4.2.9}$$

The joint density function of $(Y_1,V_1),\ldots,(Y_N,V_N)$ in the arguments $(y_1,v_1),\ldots,(y_N,v_N)$ (i.e. the likelihood function) follows from (4.2.9) by integrating out the u_n's, namely

$$\begin{aligned} L=f\left((y_1,v_1),\ldots,(y_N,v_N)\right)&=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\prod_{n=1}^{N} f_n\left(y_n,v_n,u_n\right)p_n\left(u_n|u_{n-1}\right)du_n \\ &=E_p\prod_{n=1}^{N} f_n\left(y_n,v_n,U_n\right) \end{aligned} \tag{4.2.10}$$

where E_p signifies that the expectation in the last line is taken over $U_1,\ldots,U_N$ distributed according to the joint density $p(u_1,\ldots,u_N)=p_1(u_1)\,p_2\left(u_2|u_1\right)\ldots p_N\left(u_N|u_{N-1}\right)$. Note that L is also a function of the parameters β and ψ, the GARCH parameters ν, ϕ, α_0, α_1 and β_1 (see Section 3.2.1), as well as ρ, but we do not show that explicitly for ease of notation.

Expression (4.2.10) suggests that we may use the Simulated Maximum Likelihood (SML) method to obtain an estimate for L by using a natural Monte Carlo (NMC) simulation. A NMC estimate can be obtained by drawing independent samples $\left\{u_{1,m},\ldots,u_{N,m}\right\},m=1,\ldots,M$ from $p(u_1,\ldots,u_N)$ and setting

$$\hat{L}_{NMC}=\frac{1}{M}\sum_{m=1}^{M}\prod_{n=1}^{N} f_n\left(y_n,v_n,u_{n,m}\right)\ . \tag{4.2.11}$$

However, Liesenfeld and Richard (2003) argue that such a NMC estimator is highly inefficient since it is based on a sequence of sampling densities which is directly obtained from the statistical specification of the model, ignoring critical information contained in the observed data about the underlying latent process. They also state that sampling variance increases dramatically with the sample size N and that, in practice, a prohibitively large number of Monte Carlo samples M would be required to obtain reasonably accurate estimates. This inefficiency can be dealt with by importance sampling as follows.

The likelihood function (4.2.10) can alternatively be written as

$$
\begin{aligned}
L &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\prod_{n=1}^{N} f_n\left(y_n,v_n,u_n\right)p_n\left(u_n\middle|u_{n-1}\right)du_n \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\prod_{n=1}^{N}\frac{f_n\left(y_n,v_n,u_n\right)p_n\left(u_n\middle|u_{n-1}\right)}{q_n\left(u_n\middle|u_{n-1}\right)}\prod_{n=1}^{N}q_n\left(u_n\middle|u_{n-1}\right)du_n \\
&= E_q\prod_{n=1}^{N}\frac{f_n\left(y_n,v_n,U_n\right)p_n\left(U_n\middle|U_{n-1}\right)}{q_n\left(U_n\middle|U_{n-1}\right)}\ ,
\end{aligned}
\tag{4.2.12}
$$

where $q_1\left(u\middle|s\right)=q_1\left(u\right)$ is a density, $q_2\left(u\middle|s\right),\ldots,q_N\left(u\middle|s\right)$ are transition densities and E_q signifies that the expectation in the last line is taken over $U_1,\ldots,U_N$ distributed according to the joint density $q\left(u_1,\ldots,u_N\right)=q_1\left(u_1\right)q_2\left(u_2\middle|u_1\right)\ldots q_N\left(u_N\middle|u_{N-1}\right)$, which is referred to as the importance density.

If we draw independent samples $\left\{u_{1,m},\ldots,u_{N,m}\right\},m=1,\ldots,M$ from $q\left(u_1,\ldots,u_N\right)$, we may approximate the likelihood function (4.2.12) by

$$
\hat{L}_{EIS}=\frac{1}{M}\sum_{m=1}^{M}\prod_{n=1}^{N}\frac{f_n\left(y_n,v_n,u_{n,m}\right)p_n\left(u_{n,m}\middle|u_{n-1,m}\right)}{q_n\left(u_{n,m}\middle|u_{n-1,m}\right)}\ . \tag{4.2.13}
$$

The idea of EIS is to choose the density $q\left(u_1,\ldots,u_N\right)$ for $U_1,\ldots,U_N$ such that the sampling variability of $\prod_{n=1}^{N}f_n\left(y_n,v_n,U_n\right)p_n\left(U_n\middle|U_{n-1}\right)/q_n\left(U_n\middle|U_{n-1}\right)$ is very small. This will enhance the accuracy of approximation. Liesenfeld and Richard (2003) choose

$$
q_n\left(u\middle|s\right)=\frac{p_n\left(u\middle|s\right)\varsigma_n\left(u\right)}{\chi_n\left(s\right)} \tag{4.2.14}
$$

where

$$
\varsigma_n\left(u\right)=\exp\left(a_{1,n}+a_{2,n}u+a_{3,n}u^2\right)\ \text{with}\ \left\{a_{1,n},a_{2,n},a_{3,n}\right\}\ \text{constants} \tag{4.2.15}
$$

and

$$
\chi_n\left(s\right)=\int_{-\infty}^{\infty}p_n\left(u\middle|s\right)\varsigma_n\left(u\right)du\ \text{for}\ n=1,\ldots,N\ . \tag{4.2.16}
$$

From (4.2.8), (4.2.16) and setting

$$
\sigma_n^2=\left(\left(1-\rho_n^2\right)^{-1}-2a_{3,n}\right)^{-1}\ ,\ \mu_n\left(s\right)=\sigma_t^2\left(a_{2,n}+\frac{\rho_n s}{1-\rho_n^2}\right) \tag{4.2.17}
$$

it follows that

$$\chi_n(s) = \int_{-\infty}^{\infty} p_n(u|s)\varsigma_n(u)\,du = \int_{-\infty}^{\infty} \frac{1}{\sqrt{1-\rho_n^2}}\varphi\left(\frac{u-\rho_n s}{\sqrt{1-\rho_n^2}}\right)\exp\left(a_{1,n}+a_{2,n}u+a_{3,n}u^2\right)du$$

$$= \frac{1}{\sqrt{1-\rho_n^2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{u-\rho_n s}{\sqrt{1-\rho_n^2}}\right)^2\right)\exp\left(a_{1,n}+a_{2,n}u+a_{3,n}u^2\right)du$$

$$= \frac{1}{\sqrt{1-\rho_n^2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{u^2-2u\rho_n s+\rho_n^2 s^2}{1-\rho_n^2}\right)+a_{1,n}+a_{2,n}u+a_{3,n}u^2\right)du$$

$$= \frac{1}{\sqrt{1-\rho_n^2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(u^2\left(\frac{1}{1-\rho_n^2}-2a_{3,n}\right)-2u\left(a_{2,n}+\frac{\rho_n s}{1-\rho_n^2}\right)+\frac{\rho_n^2 s^2}{1-\rho_n^2}-2a_{1,n}\right)\right)du$$

$$= \frac{1}{\sqrt{1-\rho_n^2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{u^2}{\sigma_n^2}-2u\mu_n(s)+\mu_n^2(s)-\mu_n^2(s)+\frac{\rho_n^2 s^2}{1-\rho_n^2}-2a_{1,n}\right)\right)du$$

$$= \frac{\sigma_n\exp\left(a_{1,n}+\frac{\mu_n^2(s)}{2\sigma_n^2}-\frac{\rho_n^2 s^2}{2\left(1-\rho_n^2\right)}\right)}{\sqrt{1-\rho_n^2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_n}\exp\left(-\frac{1}{2}\left(\frac{u-\mu_n(s)}{\sigma_n}\right)^2\right)du$$

$$= \frac{\sigma_n}{\sqrt{1-\rho_n^2}}\exp\left(a_{1,n}+\frac{\mu_n^2(s)}{2\sigma_n^2}-\frac{\rho_n^2 s^2}{2\left(1-\rho_n^2\right)}\right). \qquad (4.2.18)$$

Using (4.2.8), (4.2.15) and (4.2.18) we can write (4.2.14) as

$$q_n(u|s) = \frac{p_n(u|s)\varsigma_n(u)}{\chi_n(s)} = \frac{\frac{1}{\sqrt{1-\rho_n^2}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{u-\rho_n s}{\sqrt{1-\rho_n^2}}\right)^2+a_{1,n}+a_{2,n}u+a_{3,n}u^2\right)}{\frac{\sigma_n}{\sqrt{1-\rho_n^2}}\exp\left(a_{1,n}+\frac{\mu_n^2(s)}{2\sigma_n^2}-\frac{\rho_n^2 s^2}{2\left(1-\rho_n^2\right)}\right)}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_n}\exp\left(-\frac{1}{2}\left(\frac{u^2}{1-\rho_n^2}-\frac{2u\rho_n s}{1-\rho_n^2}+\frac{\rho_n^2 s^2}{1-\rho_n^2}+\frac{\mu_n^2(s)}{\sigma_n^2}-\frac{\rho_n^2 s^2}{1-\rho_n^2}-2a_{2,n}u-2a_{3,n}u^2\right)\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma_n}\exp\left(-\frac{1}{2}\left(\frac{u^2-2u\mu_n(s)+\mu_n^2(s)}{\sigma_n^2}\right)\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma_n}\exp\left(-\frac{1}{2}\left(\frac{u-\mu_n(s)}{\sigma_n}\right)^2\right). \qquad (4.2.19)$$

Thus, it follows that $q_n(u|s)$ is a $N\left(\mu_n(s),\sigma_n^2\right)$ density. Using (4.2.14) we can write

$$\prod_{n=1}^{N}\frac{f_n(y_n,v_n,u_n)\,p_n(u_n|u_{n-1})}{q_n(u_n|u_{n-1})} = \chi_1\prod_{n=1}^{N}\frac{f_n(y_n,v_n,u_n)\,\chi_{n+1}(u_n)}{\varsigma_n(u_n)} \qquad (4.2.20)$$

where we define $\chi_{N+1}(u) \equiv 1$. Liesenfeld and Richard (2003) propose choosing the constants $\{a_{1,n}, a_{2,n}, a_{3,n}\}$ to make each factor $f_n(y_n, v_n, u_n)\chi_{n+1}(u_n)/\varsigma_n(u_n)$ as nearly constant in u_n as possible using preliminary sampling and regression as follows:

- Draw independent samples $\{\tilde{u}_{1,m}, \ldots, \tilde{u}_{n,m}\}, m = 1, \ldots, M$ from $p(u_1, \ldots, u_N) = p_1(u_1)\, p_2(u_2|u_1) \cdots p_N(u_N|u_{N-1})$.
- Regress $\log \varsigma_n(\tilde{u}_{n,m}) \equiv a_{1,n} + a_{2,n}\tilde{u}_{n,m} + a_{3,n}\tilde{u}_{n,m}^2$ on $\log f_n(y_n, v_n, \tilde{u}_{n,m}) + \log \chi_{n+1}(\tilde{u}_{n,m})$, $m = 1, \ldots, \tilde{M}$ successively for $n = N, N-1, \ldots, 1$; since $\log \chi_{n+1}(\tilde{u}_{n,m})$ depends only on the previous $\{a_{1,n+1}, a_{2,n+1}, a_{3,n+1}\}$, this determines $\{a_{1,n}, a_{2,n}, a_{3,n}\}$ iteratively.
- Compute an approximation to $q(u_1, \ldots, u_N)$ using (4.2.14).
- Repeat the process a number of times, but now sample from this approximation to $q(u_1, \ldots, u_N)$, the last approximation yielding the final choice of $q(u_1, \ldots, u_N)$.

If convergence occurs in these steps, it seems reasonable to use the limiting density as $q(u_1, \ldots, u_N)$, but we also have to take computational time into account, which may dictate that we stop after a few iterations. Once we have decided on $q(u_1, \ldots, u_N)$ we may go on to the final stage where we make a final draw of M samples from $q(u_1, \ldots, u_N)$ and use (4.2.13) to approximate the likelihood function.

Note that this is a general scheme. The specific form of the function $f_n(y, v, u)$ in (4.2.8) plays no role in the formulation of the method, although it may affect its success. Liesenfeld and Richard (2003) demonstrate this method for a number of situations. We are specifically interested in using it in our present GARCH context.

When we applied this method to generated data the $\{a_{1,n}, a_{2,n}, a_{3,n}\}$'s, and thus $q(u_1, \ldots, u_N)$, converged after about six to fifteen iterations when the parameter values were within reasonable ranges. If we then applied (4.2.13) a good approximation to the likelihood function was obtained.

Variance reduction

A number of things can be done to improve the accuracy of the likelihood function estimate beyond that delivered by the EIS technique of Liesenfeld and Richard (2003). If we let

$$T = \sum_{n=1}^{N}\left(\log f_n\left(y_n, v_n, U_n\right) + \log p_n\left(U_n \middle| U_{n-1}\right) - \log q_n\left(U_n \middle| U_{n-1}\right)\right) \tag{4.2.21}$$

we can write (4.2.12) as

$$L = E_q \exp(T) \ . \tag{4.2.22}$$

Referring to (4.2.13) we may approximate (4.2.22) as follows. Define

$$t_m = \sum_{n=1}^{N}\left(\log f_n\left(y_n, v_n, u_{n,m}\right) + \log p_n\left(u_{n,m} \middle| u_{n-1,m}\right) - \log q_n\left(u_{n,m} \middle| u_{n-1,m}\right)\right) \tag{4.2.23}$$

and let $\bar{t}$ and s_t^2 be the sample mean and variance of these t_m's over m. (4.2.13) can then be written as

$$\hat{L}_{EIS} = \frac{1}{M}\sum_{m=1}^{M}\exp(t_m) \ . \tag{4.2.24}$$

In view of the central limit theorem we may expect T to be approximately normally distributed, in which case we have $L = E_q \exp(T) \simeq \exp\left(E_q T + \frac{1}{2} Var_q(T)\right)$, where Var_q signifies the variance under the joint density q given just below (4.2.12). Hence, we have an alternative estimator for L, namely

$$\hat{L}_{NOR} = \exp\left(\bar{t} + s_t^2/2\right) \ . \tag{4.2.25}$$

Inference is usually carried out in terms of $\log(L)$ which is then estimated by

$$\log\left(\hat{L}_{EIS}\right) = \log\left(1/M \sum_{m=1}^{M}\exp(t_m)\right) \text{ or } \log\left(\hat{L}_{NOR}\right) = \bar{t} + s_t^2/2 \ . \tag{4.2.26}$$

Clearly the $\hat{L}_{NOR}$ estimate has the advantage that we do not need to use the variance increasing transformation $\exp(t_m)$ used in the direct estimator $\hat{L}_{EIS}$.

Remaining variability in $\log\left(\hat{L}_{NOR}\right)$ can be reduced much by using the following control variate (see Law and Kelton (2000)) variance reduction technique. If we denote the likelihood in the case where $\rho = 0$ by L_0 we can calculate $\log(L_0)$ very accurately using the methods of the ID-BIG model of Section 3.2.3. We can also fit this $\rho = 0$ model using the EIS technique just explained.

For this purpose write $L_0 = E_q \exp(T_0)$ with T_0 as in (4.2.22), but taking $\rho = 0$ everywhere. During the simulation we generate $t_{0,m}$'s which are equivalents of (4.2.23) for the $\rho = 0$ model, calculate $\bar{t}_0$ and $s^2_{t_0}$ (the sample mean and variance of the $t_{0,m}$'s) and estimate $\log(L_0)$ by $\log(\hat{L}_0) = \bar{t}_0 + s^2_{t_0}/2$ (the equivalent of (4.2.25)).

The joint central moments of T and T_0 are given by

$$\mu_{a,b} = E_q\left(T - E_q T\right)^a E_q\left(T_0 - E_q T_0\right)^b , \tag{4.2.27}$$

which can be estimated by the corresponding sample moments

$$m_{a,b} = \frac{1}{M}\sum_{m=1}^{M}\left(t_m - \bar{t}\right)^a \left(t_{0,m} - \bar{t}_0\right)^b , \tag{4.2.28}$$

where we can set either $b = 0$ or $a = 0$ to obtain the moments of T and T_0 respectively.

The correlation between the t_m's and $t_{0,m}$'s is typically very high – of the order of 0.96 and higher – for reasonable parameter values. Figure 4.1 shows the normal Q-Q plots of and correlation between 200 t_m's and $t_{0,m}$'s sampled at the parameter values obtained by applying this method to the IBM data. The top panel illustrates that the normality assumption is reasonably valid in our example, while the bottom panel shows an empirical example of the high correlation (0.9928 in this case) between t_m and $t_{0,m}$, so that the use of these variance reduction techniques is warranted.

This suggests that we could apply the control variate technique to obtain the adjusted estimator

$$\log(\hat{L}_{NOR}) - a\left(\log(\hat{L}_0) - \log(L_0)\right), \text{ with } a = 1 , \tag{4.2.29}$$

which would be more accurate than $\log(\hat{L}_{NOR})$. This is confirmed by the empirical results below.

The optimal choice for a could be calculated, increasing the computational effort, but the extremely high correlation between t_m and $t_{0,m}$ assures that the choice of $a = 1$ will be very close to optimal.

In order to investigate the extent to which this technique reduces variance, we compare the variances of the two estimators, as follows.

We have already seen that it is reasonable to assume that T and T_0 are normally distributed, in which case we can take advantage of some well-known properties of moments of normal distributions. Specifically, it follows that

$$\mu_{3,0} = \mu_{0,3} = 0\,,\ \mu_{4,0} = 3\mu_{2,0} \text{ and } \mu_{0,4} = 3\mu_{0,2}\,. \tag{4.2.30}$$

Kendall and Stuart (1969) provide formulae for estimating the standard errors of univariate and bivariate sample moments. Using these formulae, and (4.2.30), we can write

$$\begin{aligned} &Cov\left(\bar{t}, s_t^2\right) = Cov\left(\bar{t}_0, s_{t_0}^2\right) = Cov\left(\bar{t}, s_{t_0}^2\right) = Cov\left(\bar{t}_0, s_t^2\right) = 0\,, \\ &Var\left(\hat{L}_{NOR}\right) = Var\left(\bar{t} + \frac{1}{2}s_t^2\right) = Var\left(\bar{t}\right) + \frac{1}{4}Var\left(s_t^2\right) + Cov\left(\bar{t}, s_t^2\right) \\ &\qquad = \frac{1}{M}\left(\mu_{2,0} + \frac{1}{2}\mu_{2,0}^2\right) \end{aligned} \tag{4.2.31}$$

and

$$\begin{aligned} &Var\left(\log\left(\hat{L}_{NOR}\right) - \left(\log\hat{L}_0 - \log L_0\right)\right) = Var\left(\bar{t} + \frac{1}{2}s_t^2 - \bar{t}_0 - \frac{1}{2}s_{t_0}^2\right) \\ &= Var\left(\bar{t}\right) + Var\left(\bar{t}_0\right) - 2Cov\left(\bar{t}, \bar{t}_0\right) + \frac{1}{4}\left(Var\left(s_t^2\right) + Var\left(s_{t_0}^2\right) - 2Cov\left(s_t^2, s_{t_0}^2\right)\right) \\ &\quad + Cov\left(\bar{t}, s_t^2\right) - Cov\left(\bar{t}, s_{t_0}^2\right) - Cov\left(\bar{t}_0, s_t^2\right) + Cov\left(\bar{t}_0, s_{t_0}^2\right) \\ &= \frac{1}{M}\left(\mu_{2,0} + \mu_{0,2} - 2\mu_{1,1} + \frac{1}{2}\left(\mu_{2,0}^2 + \mu_{0,2}^2 - 2\mu_{1,1}^2\right)\right) \end{aligned} \tag{4.2.32}$$

in which $\mu_{a,b}$ can be estimated by $m_{a,b}$ (see (4.2.28)).

When we use the above mentioned 200 t_m's and $t_{0,m}$'s to obtain the two log-likelihood estimates, the variance of $\log\left(\hat{L}_{NOR}\right)$, calculated using (4.2.31), is estimated as 0.03119, while the variance of $\log\left(\hat{L}_{NOR}\right) - \left(\log\left(\hat{L}_0\right) - \log\left(L_0\right)\right)$, calculated using (4.2.32) is estimated as 0.00156. This represents an almost twenty fold decrease in the variance in this example. Figure 4.2 gives an indication of the spread of t_m and $t_m - t_{0,m}$ around their means, elucidating the difference between $Var\left(\bar{t}\right)$ and $Var\left(\bar{t} - \bar{t}_0\right)$. Although the second moments s_t^2 and $s_{t_0}^2$ are not taken into account here we get a fairly good graphical representation of the difference between $Var\left(\log\left(\hat{L}_{NOR}\right)\right)$ $= Var\left(\bar{t} + s_t^2/2\right)$ and $Var\left(\log\left(\hat{L}_{NOR}\right) - \left(\log\left(\hat{L}_0\right) - \log\left(L_0\right)\right)_0\right) = Var\left(\left(\bar{t} - \bar{t}_0\right) + \left(s_t^2 - s_{t_0}^2\right)/2\right)$.

In view of these results, we will use (4.2.29) to estimate the log-likelihood function.

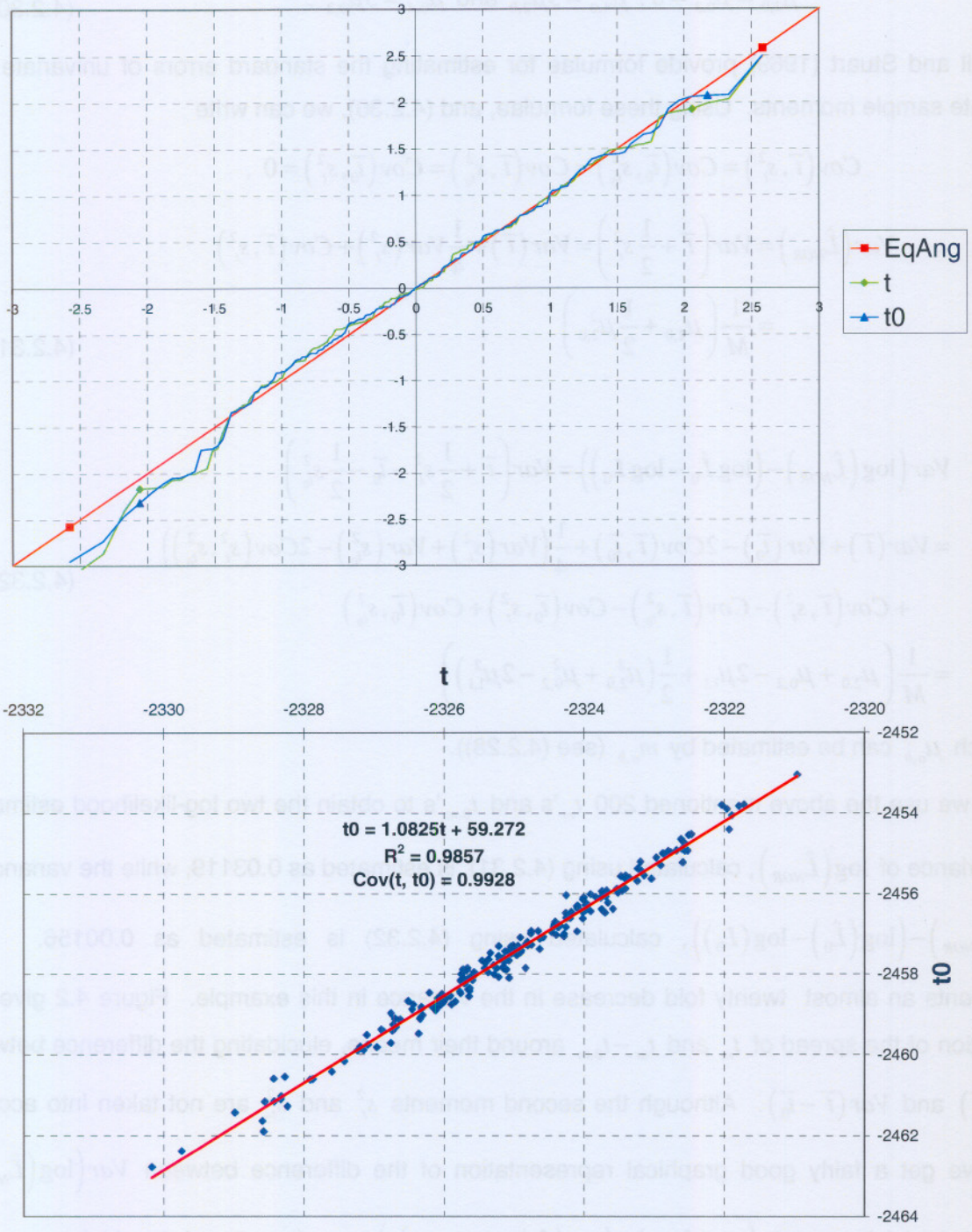

Figure 4.1: Q-Q (top panel) and scatter (bottom panel) plots of t_m and $t_{0,m}$ justifying variance reduction techniques.

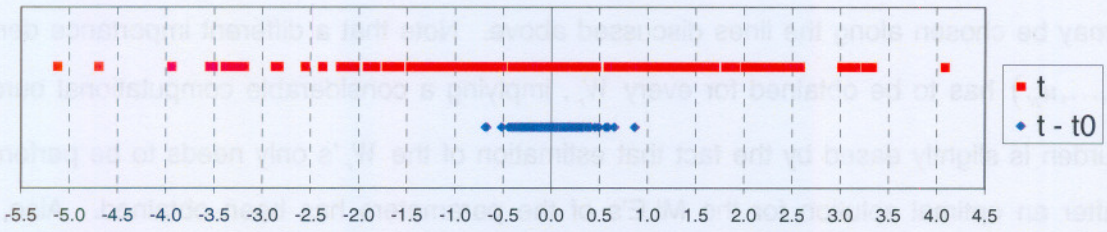

Figure 4.2: Spread of t_m and $t_m - t_{0,m}$ centred around their means.

We can fit the ID-BIG model and use the maximum likelihood estimates of the parameters as a convenient initial point from which to start searching for the maximum of the full likelihood function of the TD-ID-BIG model. Finally, it is important to use common random numbers when calculating the likelihood function at different parameter values during the maximising process, since each time the likelihood function is calculated, it should be dependent on the parameter values only, not on the specific random numbers used. We achieve this by setting up a fixed matrix of independent standard normal random numbers and each time computing the $\tilde{u}_{n,m}$'s by simple transformations from the entries in this matrix.

We can also extend the EIS estimation technique to handle other inference issues, for instance estimation of the random impact factor W_r for day r. If we define

$$\tilde{f}_n^{(r)}(y_n, v_n, u_n) = \begin{cases} g(u_n) f_n(y_n, v_n, u_n) & \text{for } n = r \\ f_n(y_n, v_n, u_n) & \text{for } n \neq r \end{cases} \tag{4.2.33}$$

and use (4.2.3) and (4.2.10) we have

$$\begin{aligned} E\left(W_r \middle| (y_1, v_1), \ldots, (y_N, v_N)\right) &= E\left(g(U_n) \middle| (y_1, v_1), \ldots, (y_N, v_N)\right) \\ &= \int_{-\infty}^{\infty} g(u_r) f\left(u_r \middle| (y_1, v_1), \ldots, (y_N, v_N)\right) du_r \\ &= \int_{-\infty}^{\infty} g(u_r) \frac{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f\left((y_1, v_1, u_1), \ldots, (y_N, v_N, u_N)\right) \prod_{n=1; n \neq r}^{N} du_n}{f\left((y_1, v_1), \ldots, (y_N, v_N)\right)} du_r \\ &= \frac{1}{L} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(u_r) \prod_{n=1}^{N} f_n(y_n, v_n, u_n) p_n\left(u_n \middle| u_{n-1}\right) du_n \\ &= \frac{1}{L} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{n=1}^{N} \frac{\tilde{f}_n^{(r)}(y_n, v_n, u_n) p_n\left(u_n \middle| u_{n-1}\right)}{\tilde{q}_n^{(r)}\left(u_n \middle| u_{n-1}\right)} \prod_{n=1}^{N} \tilde{q}_n^{(r)}\left(u_n \middle| u_{n-1}\right) du_n \\ &= \frac{1}{L} E_{\tilde{q}^{(r)}} \prod_{n=1}^{N} \frac{\tilde{f}_n^{(r)}(y_n, v_n, U_n) p_n\left(U_n \middle| U_{n-1}\right)}{\tilde{q}_n^{(r)}\left(U_n \middle| U_{n-1}\right)} \end{aligned} \tag{4.2.34}$$

where $\tilde{q}^{(r)}(u_1,\ldots,u_N)=\tilde{q}_1^{(r)}(u_1)\tilde{q}_2^{(r)}(u_2|u_1)\cdots\tilde{q}_N^{(r)}(u_N|u_{N-1})$ is a further relevant importance density which may be chosen along the lines discussed above. Note that a different importance density $\tilde{q}^{(r)}(u_1,\ldots,u_N)$ has to be obtained for every W_r, implying a considerable computational burden. This burden is slightly eased by the fact that estimation of the W_r's only needs to be performed once after an optimal solution for the MLE's of the parameters has been obtained. Also, the constants $\{a_{1,n}^{(r)},a_{2,n}^{(r)},a_{3,n}^{(r)}\}$ that define $\tilde{q}^{(r)}(u_1,\ldots,u_N)$ are all reasonably close to the constants $\{a_{1,n},a_{2,n},a_{3,n}\}$ that define $q(u_1,\ldots,u_N)$. Thus, a smaller number of iterations are needed to calculate each $\tilde{q}^{(r)}(u_1,\ldots,u_N)$ when we start with the constants $\{a_{1,n},a_{2,n},a_{3,n}\}$ obtained from calculating the final $q(u_1,\ldots,u_N)$. Once we have $\tilde{q}^{(r)}(u_1,\ldots,u_N)$ we can introduce $\tilde{t}_m^{(r)}$'s analogous to the t_m's in (4.2.23) and use them to calculate

$$\hat{W}_{EIS}^{(r)}=\frac{1}{\hat{L}_{EIS}}\frac{1}{M}\sum_{m=1}^{M}\exp\left(\tilde{t}_m^{(r)}\right). \qquad (4.2.35)$$

We may also apply the variance reduction techniques as above to obtain more accurate estimates, leading to the analogous estimate $\hat{W}_{NOR}^{(r)}$. Which ever way we estimate the $W_n^{(r)}$'s we can check whether they conform to the $AR(\rho)-UIG(\psi)$ process and these estimates, together with the residual analysis techniques described for the ID-BIG model of Section 3.2.3, may then be used to perform residual analysis.

Extending the D-NIG and HLC-BIG models

Suppose we have no intraday data available, only daily returns as in Section 3.2.2, and we want to allow for lag1 autocorrelation in W_n. Thus, we want to extend the D-NIG model of Section 3.2.2 to a time dependent TD-D-NIG model. We can follow the same EIS method presented above (except that V_n does not enter the model) and carry out the same steps to maximise the likelihood function for this daily returns case, using the maximum likelihood estimates obtained through fitting the D-NIG model as a starting point. However, since it is actually the intraday data that allows us to estimate the W_n's accurately (compare, for instance, Figures 3.29 and 3.32), we can expect to again obtain poor estimates for W_n and, consequently, very poor estimates for the autocorrelation

in W_n. We implemented this method for the daily returns also and found that this is indeed the case. It is virtually impossible to estimate the W_n's (and, consequently, ρ) accurately using only daily data (see also the arguments in Section 4.3.1 below). Since a model allowing for lag1 autocorrelation in W_n while incorporating only daily data is doomed from the outset, we do not consider it further. We could also follow the EIS method presented above to extend the HLC-BIG of Section 3.2.4 to a time dependent TD-HLC-BIG model for when we have high and low prices available in addition to daily close prices. We have seen in Section 2.7.5 that the information gain through having available the high and low prices corresponds to the information gain in the intraday data situation with about 8 periods per trading day (since $t = I - 1 = 7$ with I the number of periods). However, in Section 4.3.1 below we argue that results obtained through fitting the TD-ID-BIG on intraday data should not be considered fully reliable when data for 6 intraday periods are available, and may only be considered reasonably reliable when data for 13 intraday periods are available. Since using 6 or 8 intraday periods data would lead to very similar results, we do not expect the results of such a potential TD-HLC-BIG model to be reliable, and we do not pursue this matter further.

4.3 Empirical Illustrations of Time Dependent Models

We shall analyse the following data sets in this section:

- GenTD-BIG: We generate the same GenBIG data set (with the same underlying $Z_{t,n}$ processes) as in Section 3.3, with the exception that the W_n's are now generated from an $AR(\rho)-UIG(\psi)$ process with $\rho = 0.5$ (again suggested by typical empirical values), instead of being iid $UIG(\psi)$-distributed. In addition to the 5, 10 and 30 minute interval data, we also generate 65 minute interval data, which corresponds to 6 intervals per day, to show that there is a rapid decline in the accuracy of the estimate for ρ as intraday data frequency decreases.
- IBM: We use the same intraday IBM data as in Section 3.3.

Again, we use the generated data sets to illustrate the theory under known parameter values and as a test for our programs and derivations. We still fit AR(1)-GARCH(1,1) models and we provide results similar to those of Section 3.3 for the models presented above in terms of these data sets.

4.3.1 GenTD-BIG Data

We will now present the results obtained when fitting the TD-ID-BIG model of Section 4.2 to the GenTD-BIG data. For comparative purposes, we also show the maximum likelihood (ML) estimates of the parameters obtained from fitting the ID-BIG model of Section 3.2.3 to the GenTD-BIG data. Since the data was generated using known parameter values, this section serves as a test for our programs, an illustration of the underlying theory and an inspection of the changes in accuracy of the parameter values as more and more data is used.

ML estimates

Table 4.1 gives the ML estimates of the parameters, with standard errors in brackets, of the ID-BIG and TD-ID-BIG model fits to the 5,10,30 and 65 minute GenTD-BIG data. We consider the TD-ID-BIG results first and compare them with the results of the ID-BIG model in the next paragraph. We expect the TD-ID-BIG model to fit quite well and, taking standard errors into account, we see that this is indeed the case for the 5, 10 and 30 minute data. Note that the standard errors generally become smaller as more data is used, again illustrating the accuracy gain as more data becomes available. The results of the model fit to the 65 minute data serves to illustrate the drastic rate at which some of the parameter estimates become less accurate when fewer and fewer intraday periods (6 in this case) are available. The estimates $\hat{\psi}$, $\hat{\alpha}_0$ and $\hat{\rho}$ are most susceptible to inaccuracy. For instance, the standard error of $\hat{\rho}$ changes from 0.0322 in the 5 minute case to 0.0506 in the 30 minute case, which amounts to a 57.1% increase in standard error while the number of intervals decreased sixfold from 78 to 13. If the number of intervals decreases further from 13 to 6 (just over twofold), the standard error changes to 0.0727, which amounts to a change of 43.8% on the 30 minute standard error. This illustrates the exponential growth in the standard error of $\hat{\rho}$ as fewer daily intervals are available. This inaccuracy explodes in the limiting case where only daily data is available.

Parameter	GenTD-BIG	ID-BIG65	ID-BIG30	ID-BIG10	ID-BIG5
β	0.20	0.1875 (0.0673)	0.1723 (0.0567)	0.1749 (0.0497)	0.1829 (0.0490)
ψ	2.00	1.7494 (0.0977)	1.9459 (0.0876)	1.9619 (0.0564)	1.9586 (0.0522)
ν	0.00	0.0009 (0.0006)	0.0010 (0.0005)	0.0009 (0.0004)	0.0008 (0.0004)
ϕ	0.15	0.1100 (0.0307)	0.1336 (0.0286)	0.1413 (0.0251)	0.1388 (0.0244)
$\alpha_0 x 10^5$	1.00	0.5788 (0.1089)	0.9952 (0.1512)	1.1024 (0.1410)	1.1756 (0.1465)
α_1	0.10	0.1081 (0.0112)	0.1412 (0.0136)	0.1532 (0.0129)	0.1560 (0.0128)
β_1	0.80	0.8026 (0.0196)	0.7542 (0.0242)	0.7441 (0.0216)	0.7363 (0.0222)
ρ	0.50	0.1625 (0.0327)	0.2433 (0.0309)	0.3570 (0.0297)	0.4338 (0.0287)
Parameter	**GenTD-BIG**	**TD-ID-BIG65**	**TD-ID-BIG30**	**TD-ID-BIG10**	**TD-ID-BIG5**
β	0.20	0.1842 (0.0624)	0.1768 (0.0572)	0.1861 (0.0532)	0.1977 (0.0528)
ψ	2.00	1.7050 (0.1005)	1.9224 (0.0873)	1.9333 (0.0689)	1.9232 (0.0671)
ν	0.00	0.0009 (0.0005)	0.0009 (0.0005)	0.0008 (0.0004)	0.0007 (0.0004)
ϕ	0.15	0.1062 (0.0286)	0.1323 (0.0286)	0.1350 (0.0282)	0.1323 (0.0282)
$\alpha_0 x 10^5$	1.00	0.5266 (0.1206)	0.8314 (0.1604)	0.8875 (0.1596)	0.9216 (0.1732)
α_1	0.10	0.0747 (0.0127)	0.1050 (0.0133)	0.1096 (0.0122)	0.1053 (0.0118)
β_1	0.80	0.8424 (0.0222)	0.8056 (0.0238)	0.8043 (0.0217)	0.8063 (0.0223)
ρ	0.50	0.4443 (0.0727)	0.4466 (0.0506)	0.4721 (0.0366)	0.5205 (0.0322)

Table 4.1: ML estimates of ID-BIG and TD-ID-BIG fits to GenTD-BIG data.

Comparing the ID-BIG and TD-ID-BIG ML estimates

Recall that the GenTD-BIG data was generated with dependence in the random impact factors W_n incorporated into the generating process. Since the ID-BIG model makes no provision for such dependence, we expect the ID-BIG model fits to the GenTD-BIG data to be inferior to the TD-ID-BIG model fits. At the end of Chapter 3 we asked the question of to what extent the estimates obtained from fitting the ID-BIG model to data with dependence in the W_n's will be influenced by this dependence. We will now shed some light on this topic. Table 4.1 also shows the ML estimates and standard errors obtained from fitting the ID-BIG model to the GenTD-BIG data. The estimate for ρ was obtained from fitting an $AR(1)$ process to the $\log \hat{W}_{XV,n}$ estimates, since (A.2.10) in the Appendix shows that the autocorrelation in the $\log \hat{W}_{XV,n}$ series should be approximately equal to ρ. For the ID-BIG fit to the 5 minute data we see that the estimates of α_1, β_1 and ρ are more than two standard errors from the true values used to generate the data. For the 10, 30 and 65 minute data the situation gets worse as lower frequency data is used, with the estimates of α_1, β_1 and ρ being the least accurate. Further empirical studies have shown that,

although the estimates of β, ν and ϕ are not affected by dependence in $\{W_n\}$, the estimates of α_1, β_1 and ρ (and to a lesser extent α_0 and ψ) are indeed severely biased by this dependence.

Quality of fit

As to checking the quality of the TD-ID-BIG fits, Figures 4.3 to 4.10 show the PIT and Q-Q plots for the residuals $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ as well as the estimates $\hat{W}_{NOR,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ for the TD-ID-BIG model fits to the 5, 10, 30 and 65 minute GenTD-BIG data. We see that the distributions of the residuals and estimates of the 5 and 10 minute models correspond very well with their model implied distributions so that, as expected, the model fits the 5 and 10 minute generated data well. However, as we expect, the quality of the fit degrades as less data is utilised. This can be seen from the larger discrepancy between the equi-angular line and the PIT and Q-Q plots for fits to the 30 and 65 minute data.

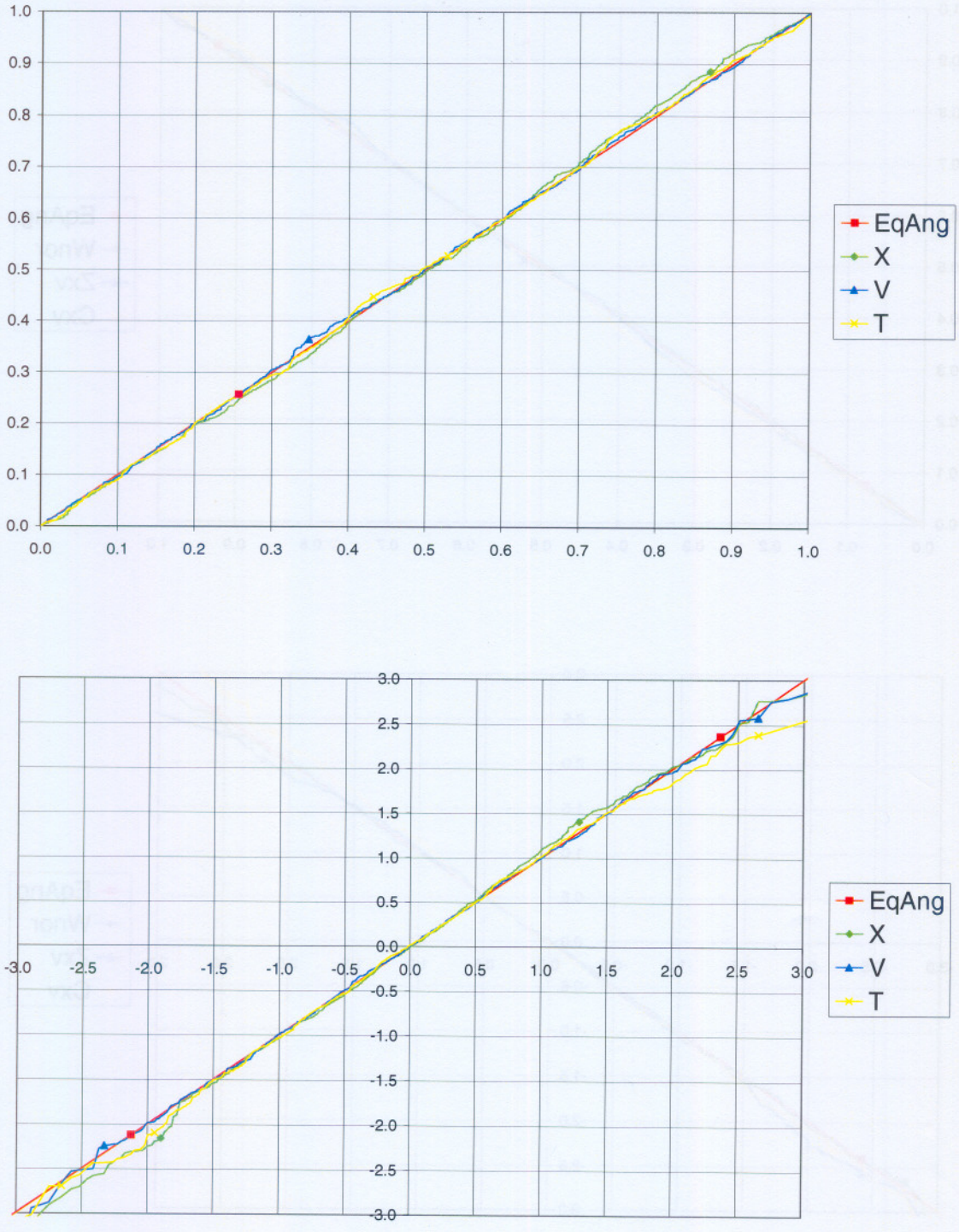

Figure 4.3: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking TD-ID-BIG5 fit to GenTD-BIG data.

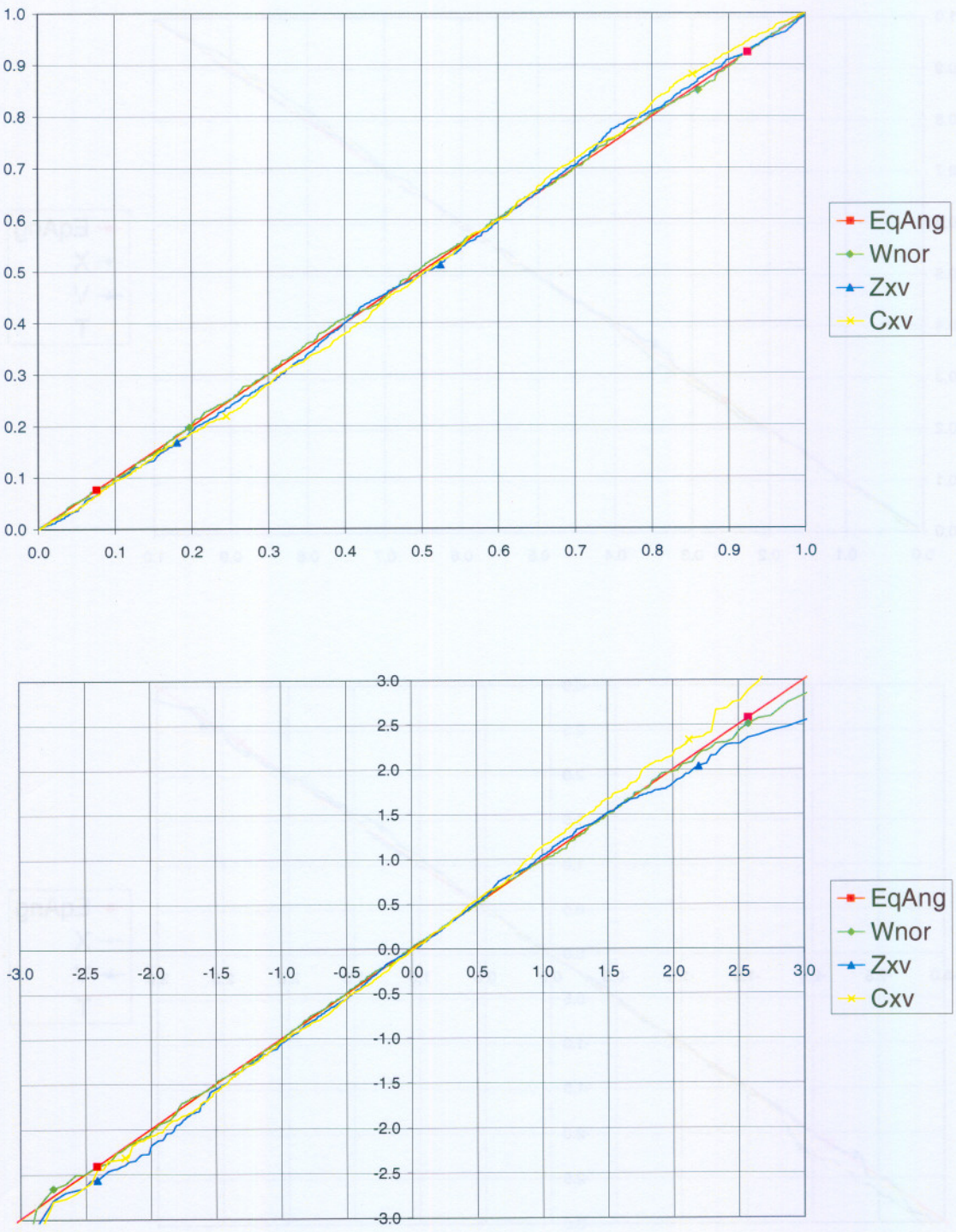

Figure 4.4: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{NOR,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking TD-ID-BIG5 fit to GenTD-BIG data.

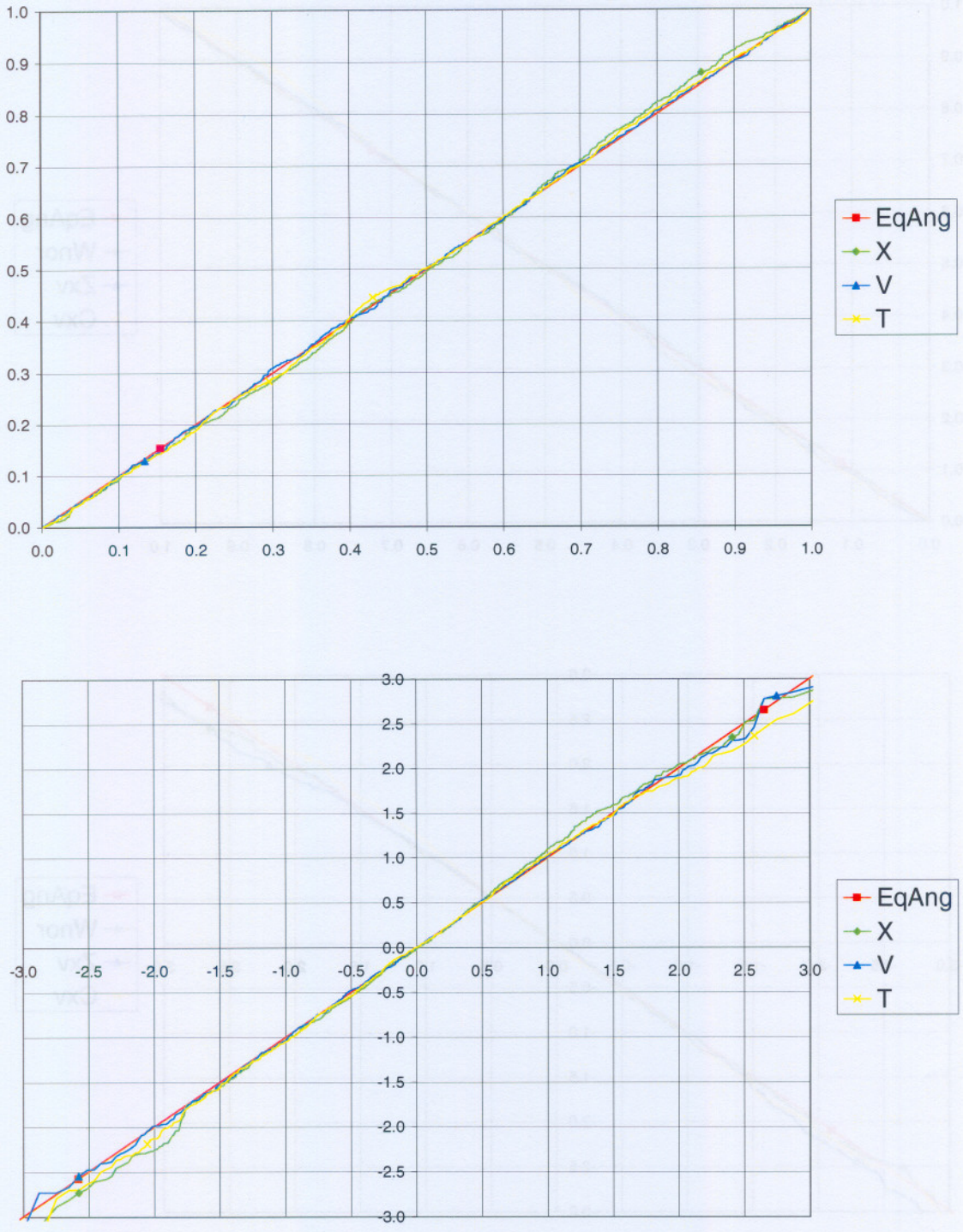

Figure 4.5: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking TD-ID-BIG10 fit to GenTD-BIG data.

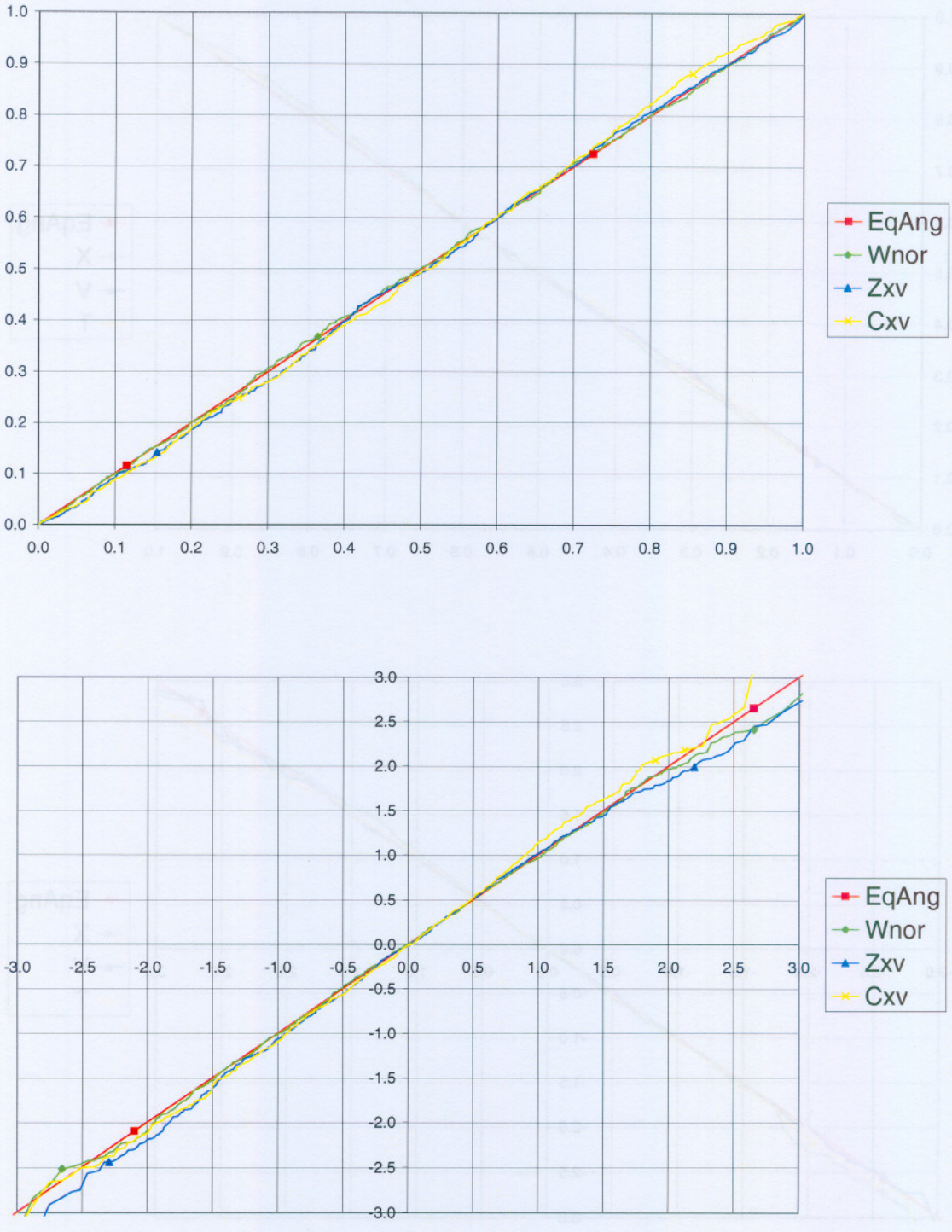

Figure 4.6: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{NOR,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking TD-ID-BIG10 fit to GenTD-BIG data.

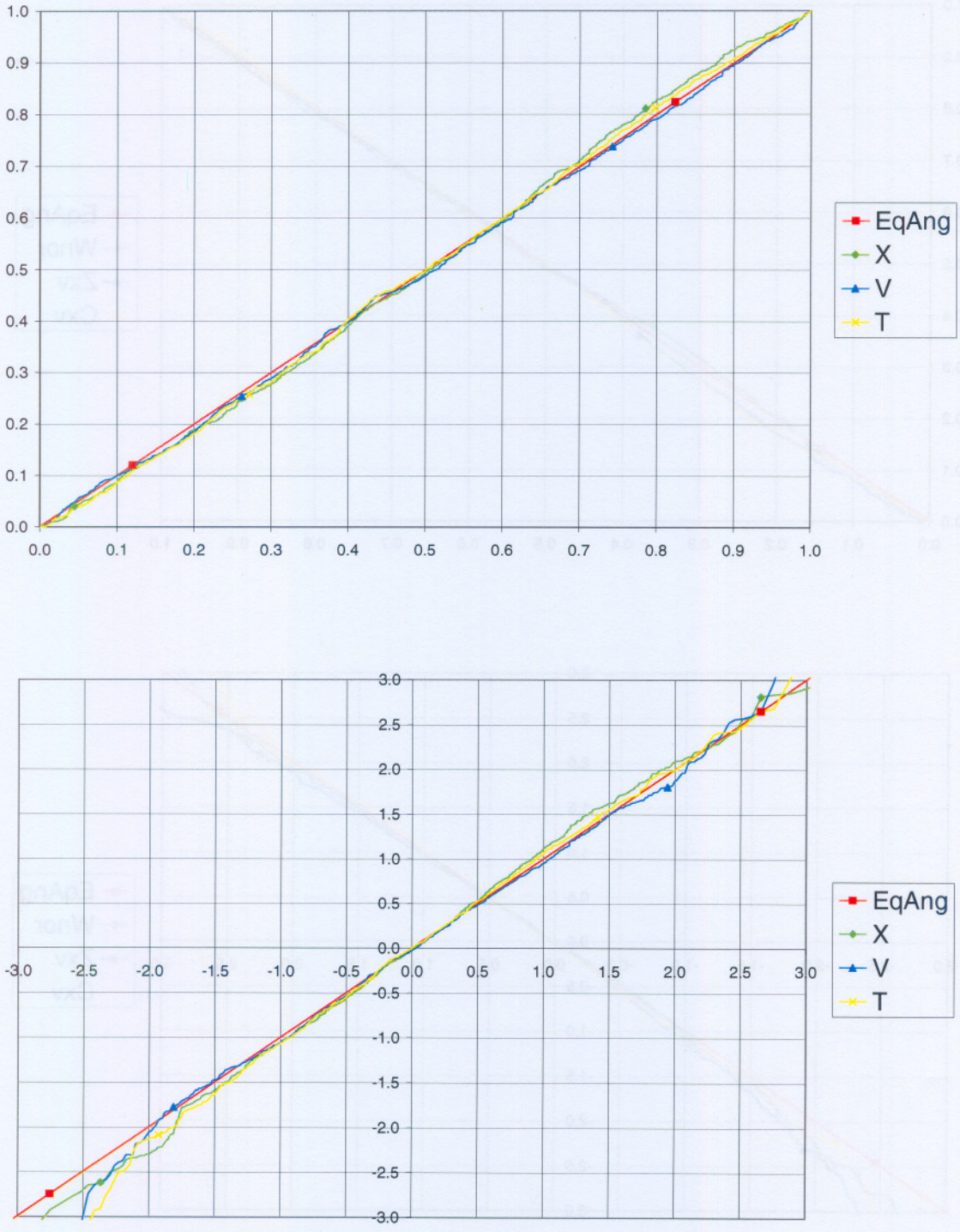

Figure 4.7: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking TD-ID-BIG30 fit to GenTD-BIG data.

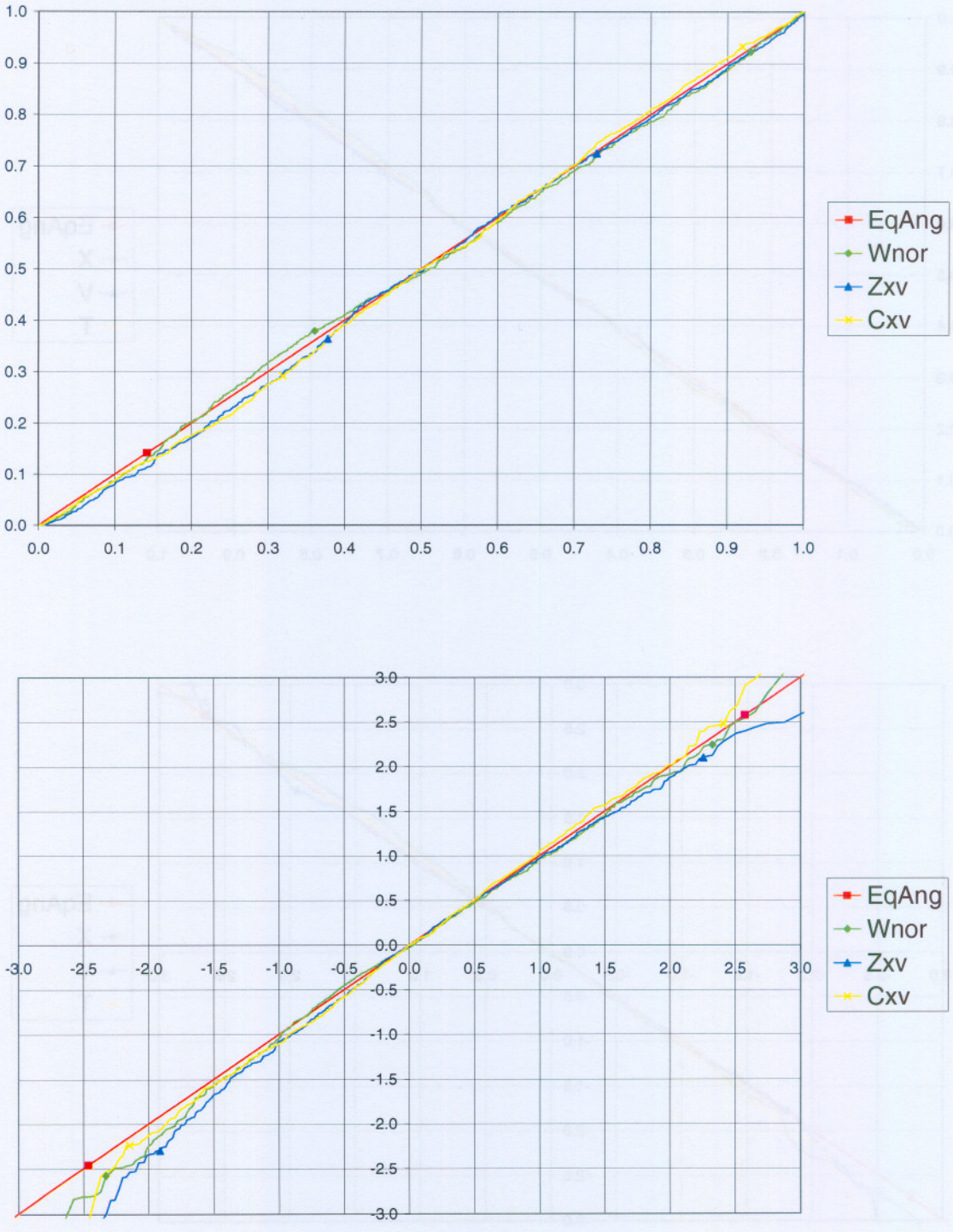

Figure 4.8: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{NOR,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking TD-ID-BIG30 fit to GenTD-BIG data.

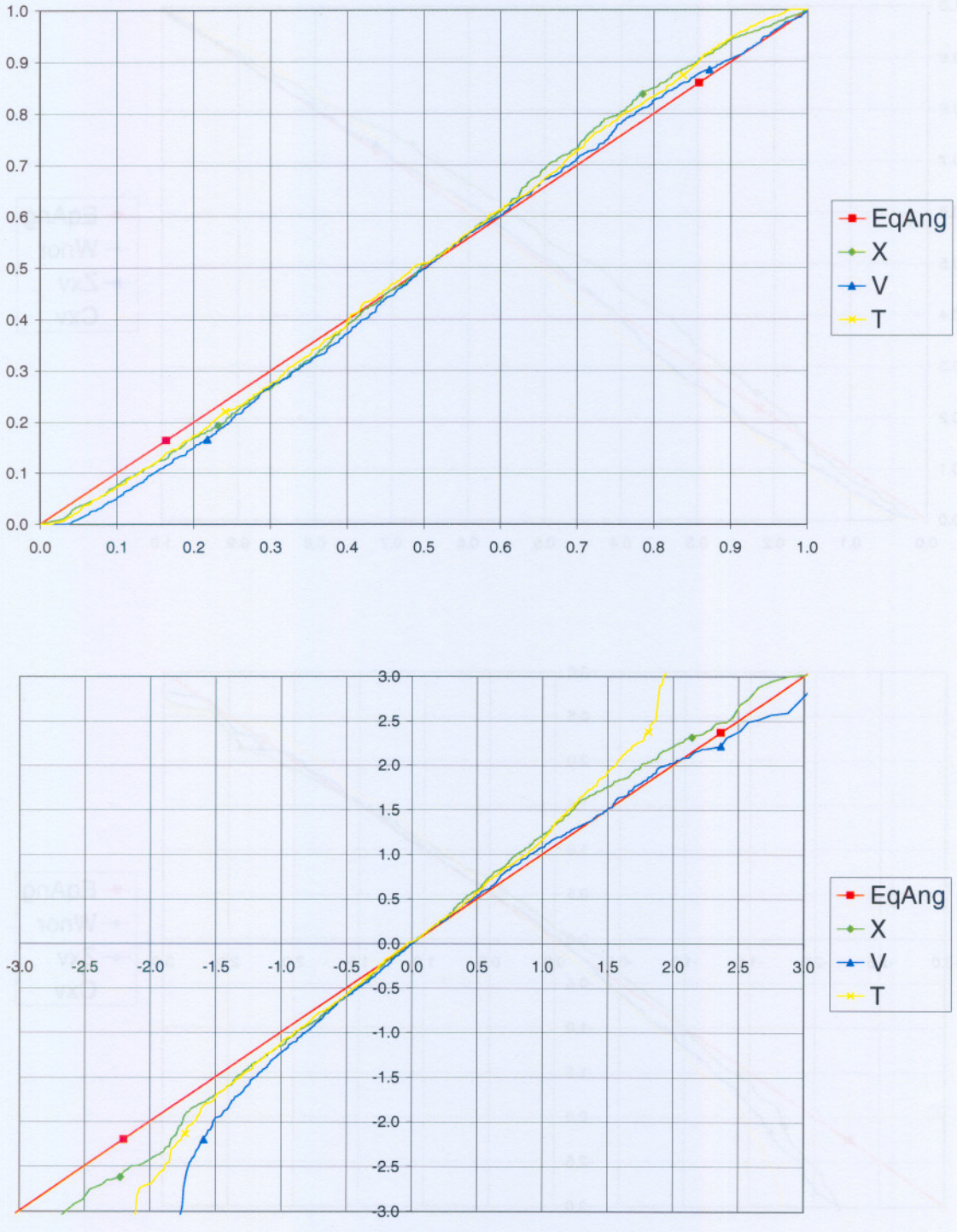

Figure 4.9: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking TD-ID-BIG65 fit to GenTD-BIG data.

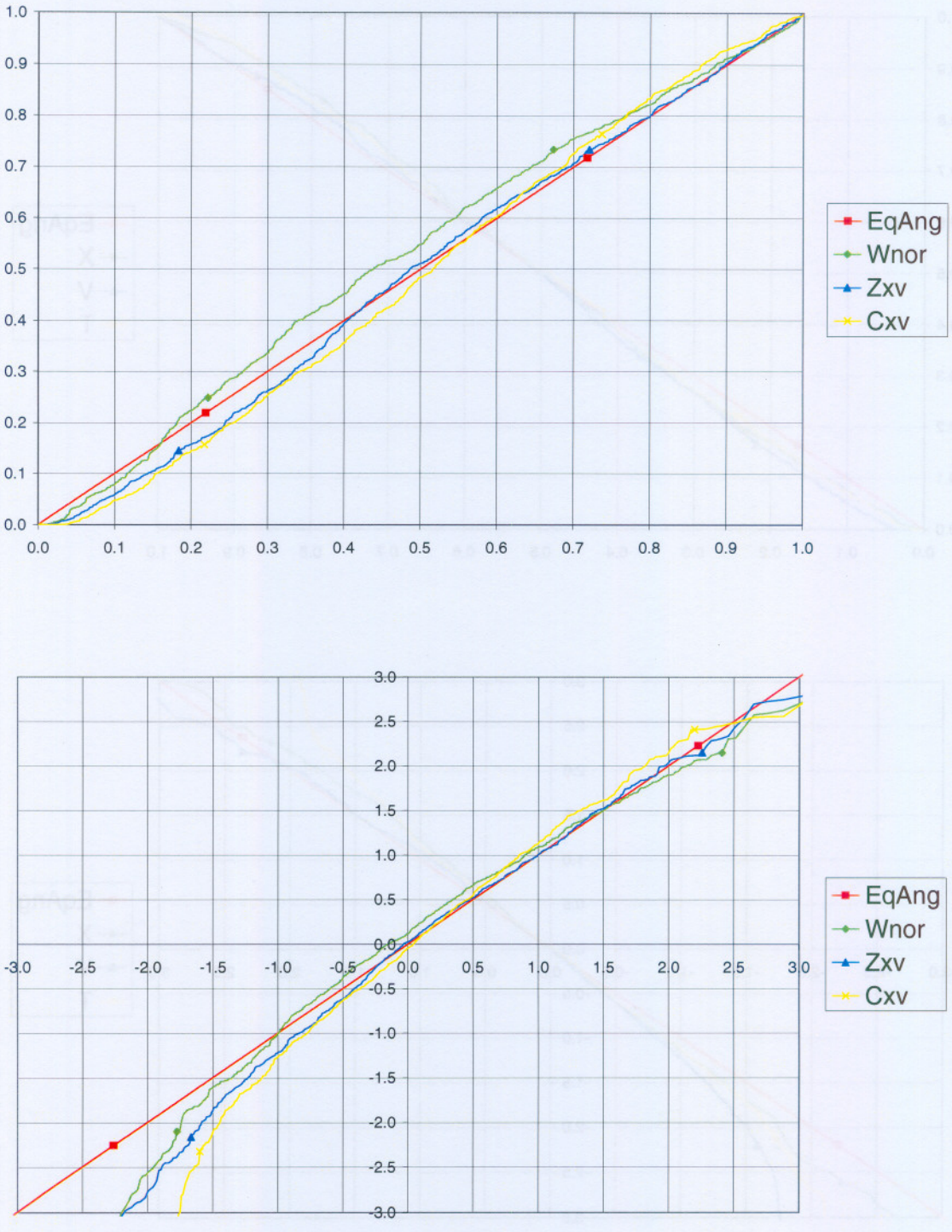

Figure 4.10: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{NOR,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking TD-ID-BIG65 fit to GenTD-BIG data.

Volatility estimates

When we examine the random impact factors for the TD-ID-BIG fits, we see similar results and trends as in previous sections. Figures 4.11 to 4.14 show W_n, $\hat{h}_n$, $\hat{W}_{NOR,n}$, $\hat{h}_n\hat{W}_{NOR,n}$ and R_n for the TD-ID-BIG fits to 5, 10, 30 and 65 minute data. Again, note that the true values for the random impact factors W_n are not known when using real-world data. We show W_n here to evaluate the goodness of the fits when using generated data. It is clear that $\hat{W}_{NOR,n}$ estimates W_n extremely accurately in the 5 minute data case and that the accuracy wanes slightly as less-frequent intervals are used. We also see how the ordinary GARCH volatility estimator $\hat{h}_n$ is adjusted by $\hat{h}_n\hat{W}_{NOR,n}$, to become much closer to the realised volatility R_n. Again, we clearly see that $\hat{h}_n$ is a smoothed volatility estimator. Figures 4.15 to 4.18 plot $\hat{h}_n$ and $\hat{h}_n\hat{W}_{NOR,n}$ against R_n for all the TD-ID-BIG fits to the GenTD-BIG data. This also shows that $\hat{h}_n\hat{W}_{NOR,n}$ tracks R_n much closer than $\hat{h}_n$ does, and that the difference between $\hat{h}_n\hat{W}_{NOR,n}$ and R_n becomes greater when fewer daily intervals are used.

Testing for independence

Finally, we consider the results of the DW, Q and LM tests on the residuals and estimates obtained from the TD-ID-BIG fits to 5, 10, 30 and 65 minute GenTD-BIG data, as given in Tables 4.2 to 4.5. The Q and LM tests indicate no remaining ARCH effects for any of the variables (after allowing for lag1 autocorrelation where appropriate), with the following two exceptions on $\hat{X}_n$. Firstly, the Q and LM tests give indications of possible second order ARCH effects for the fit to the 5 minute data. This occurrence is not too disturbing since at 95% confidence one failure in twenty tests is to be expected. Secondly, and more problematic, the Q and LM tests indicate remaining second to sixth order ARCH effects for the fit to the 65 minute data. This is further evidence that the results of the TD-ID-BIG fit to 65 minute interval data should be considered unreliable. As expected, the DW test strongly indicates lag1 autocorrelation in the $\log\hat{W}_{NOR,n}$, $\log\hat{V}_n$ and $\log\hat{C}_{XV,n}$ estimates and residuals, for all the different fits. For instance, the lag1 autocorrelation in $\log\hat{W}_{NOR,n}$ is 0.568 for the 5 minute data and 0.550 for the 10 minute data. Taking standard errors into account these values correspond satisfactorily to the $\hat{\rho}$ estimates of 0.5205 and 0.4721 respectively obtained from the model fits. After allowing for the lag1 autocorrelation, the DW test indicates no further autocorrelation in the estimates and residuals for the 5 and 10 minute results. However, there are indications of further positive autocorrelation in $\log\hat{W}_{NOR,n}$ for the 30 and 65 minute results.

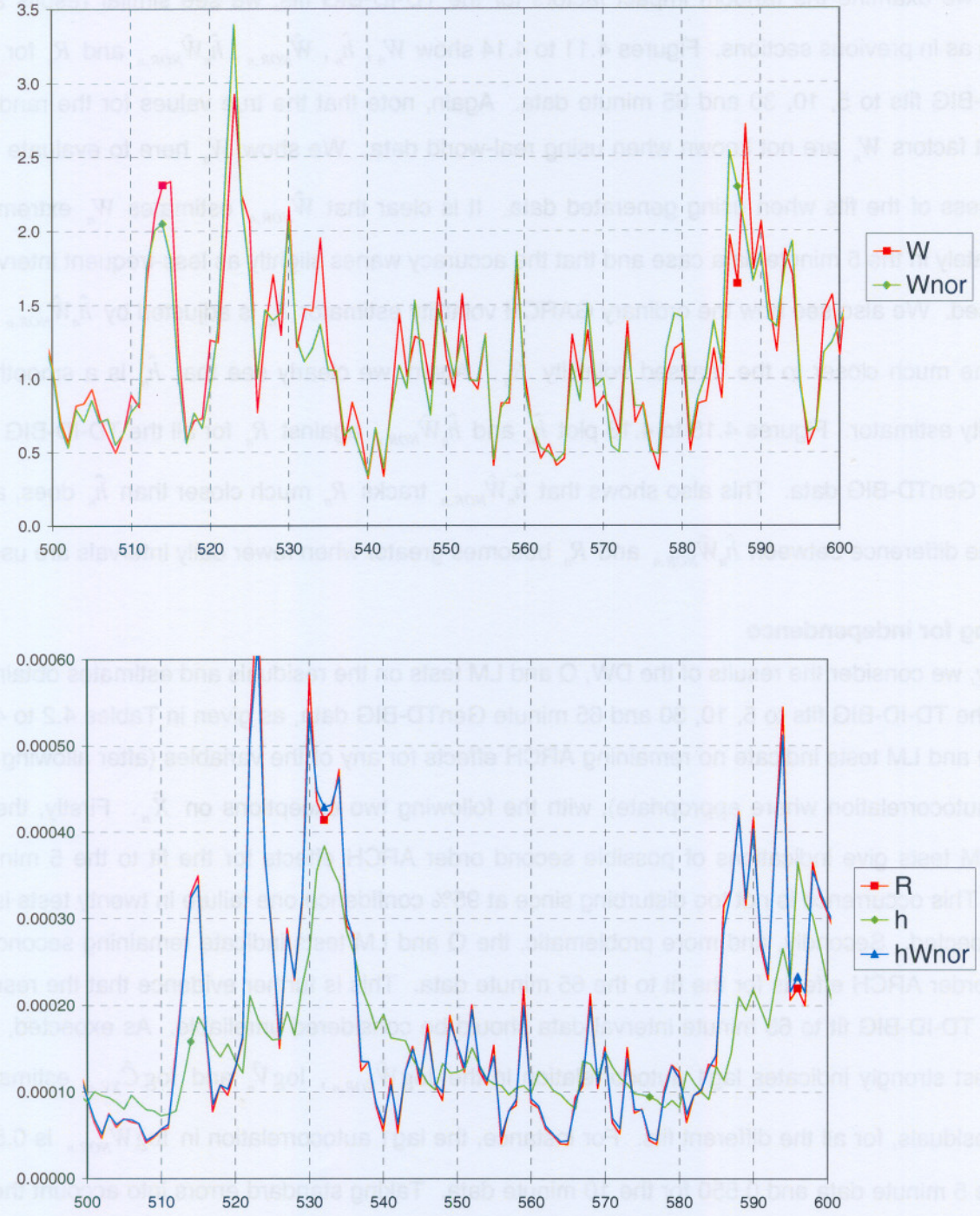

Figure 4.11: Plots of W_n and $\hat{W}_{NOR,n}$ (top panel) and R_n, $\hat{h}_n$ and $\hat{h}_n\hat{W}_{NOR,n}$ (bottom panel) for TD-ID-BIG5 fit to GenTD-BIG data.

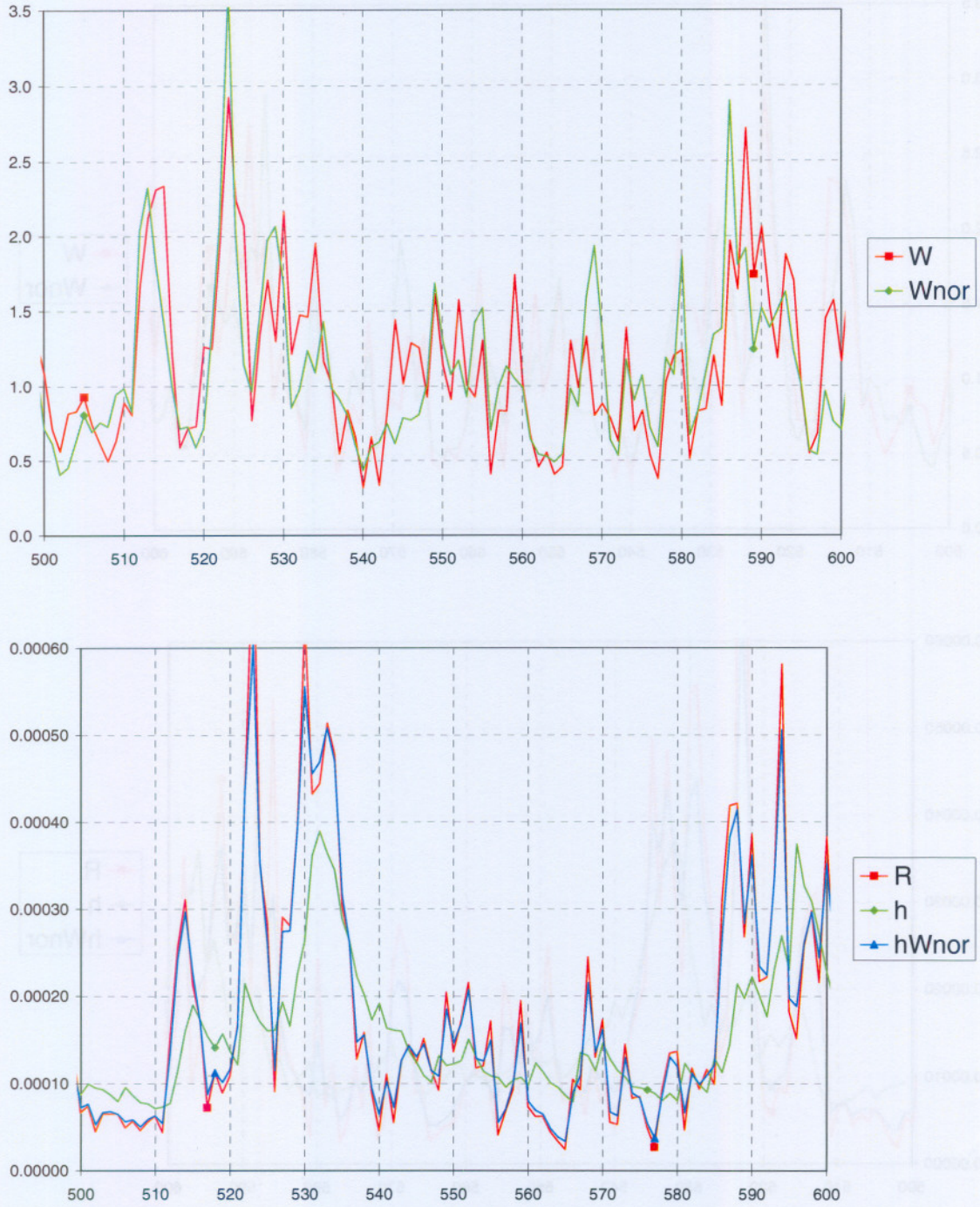

Figure 4.12: Plots of W_n and $\hat{W}_{NOR,n}$ (top panel) and R_n, $\hat{h}_n$ and $\hat{h}_n\hat{W}_{NOR,n}$ (bottom panel) for TD-ID-BIG10 fit to GenTD-BIG data.

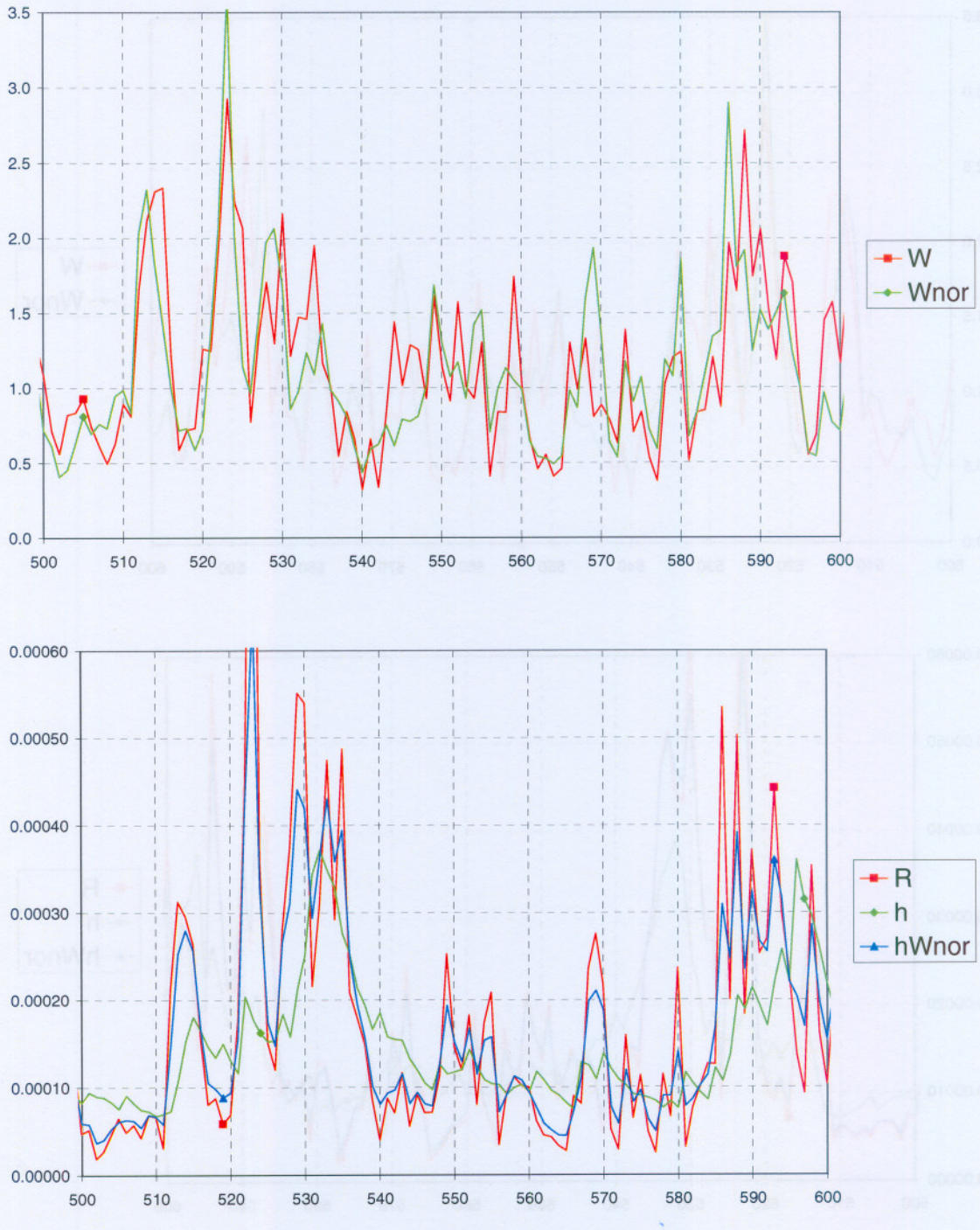

Figure 4.13: Plots of W_n and $\hat{W}_{NOR,n}$ (top panel) and R_n, $\hat{h}_n$ and $\hat{h}_n\hat{W}_{NOR,n}$ (bottom panel) for TD-ID-BIG30 fit to GenTD-BIG data.

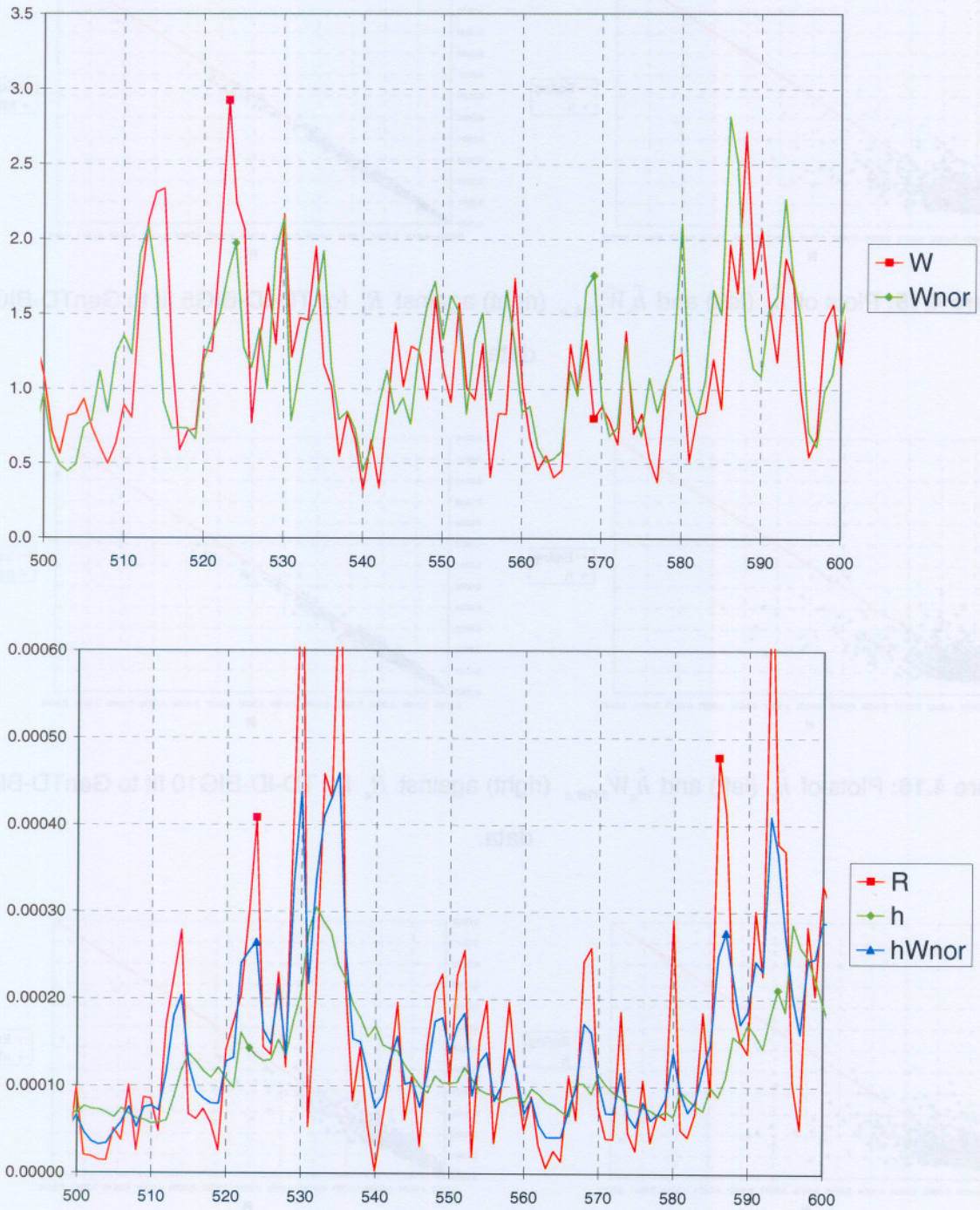

Figure 4.14: Plots of W_n and $\hat{W}_{NOR,n}$ (top panel) and R_n, $\hat{h}_n$ and $\hat{h}_n\hat{W}_{NOR,n}$ (bottom panel) for TD-ID-BIG65 fit to GenTD-BIG data.

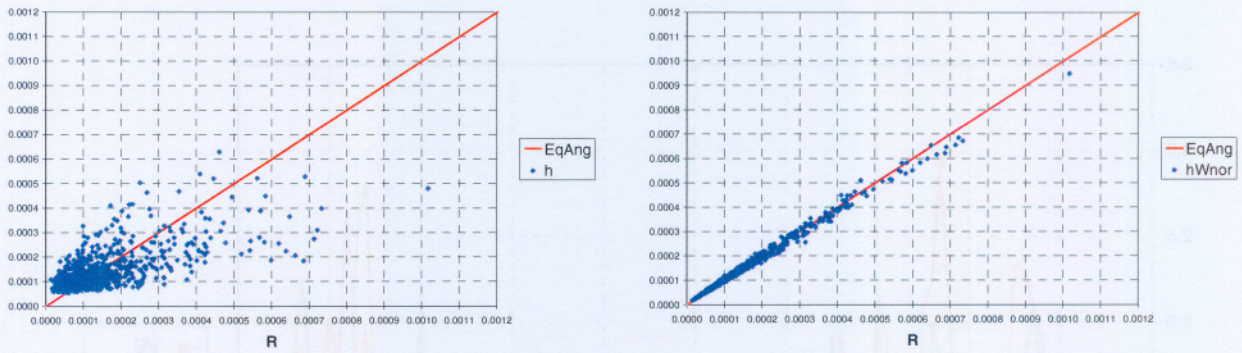

Figure 4.15: Plots of $\hat{h}_n$ (left) and $\hat{h}_n\hat{W}_{NOR,n}$ (right) against R_n for TD-ID-BIG5 fit to GenTD-BIG data.

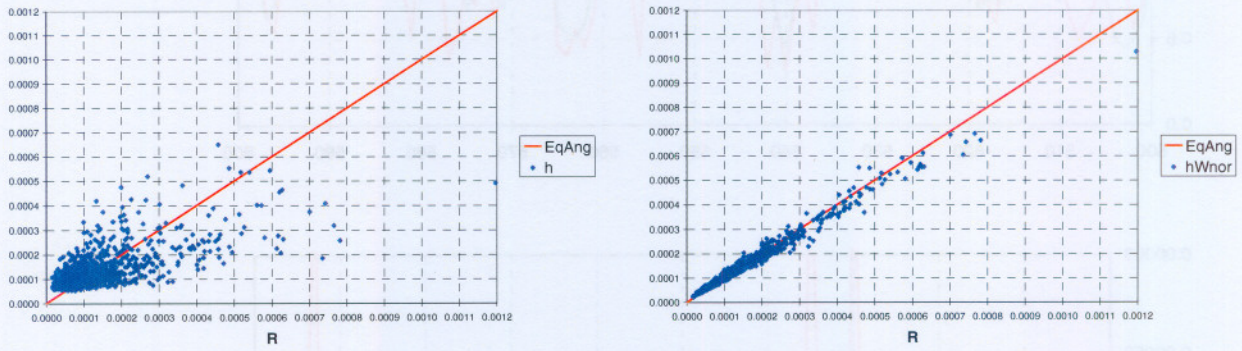

Figure 4.16: Plots of $\hat{h}_n$ (left) and $\hat{h}_n\hat{W}_{NOR,n}$ (right) against R_n for TD-ID-BIG10 fit to GenTD-BIG data.

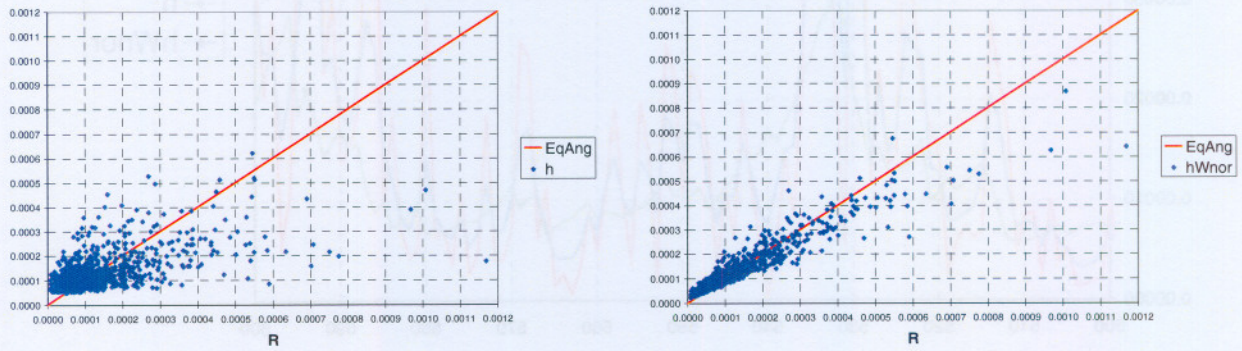

Figure 4.17: Plots of $\hat{h}_n$ (left) and $\hat{h}_n\hat{W}_{NOR,n}$ (right) against R_n for TD-ID-BIG30 fit to GenTD-BIG data.

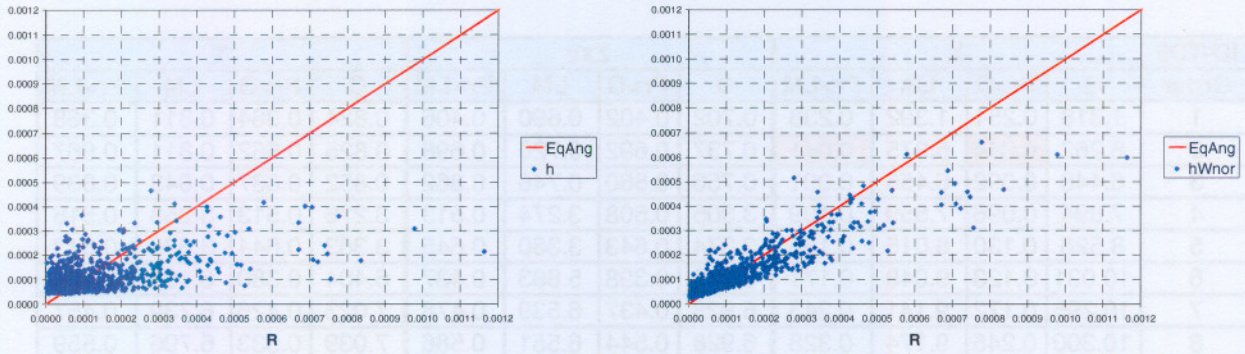

Figure 4.18: Plots of $\hat{h}_n$ (left) and $\hat{h}_n\hat{W}_{NOR,n}$ (right) against R_n for TD-ID-BIG65 fit to GenTD-BIG data.

The results presented above indicate that it would be sensible to obtain as much intraday data as possible when fitting the TD-ID-BIG model (it seems that the results of the fit to 30 minute data may still be used with reasonable confidence), while bearing in mind the possible inaccuracies of the fit when inadequate intraday data is available (the results of the fit to 65 minute data must be considered not fully reliable).

ID-TD5	X				Zxv				T			
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	1.318	0.251	1.392	0.238	0.702	0.402	0.690	0.406	0.824	0.364	0.811	0.368
2	6.262	0.044	6.375	0.041	0.737	0.692	0.720	0.698	0.825	0.662	0.811	0.667
3	6.448	0.092	6.456	0.091	0.756	0.860	0.746	0.862	0.852	0.837	0.843	0.839
4	7.894	0.096	7.559	0.109	3.305	0.508	3.274	0.513	3.276	0.513	3.258	0.516
5	8.526	0.130	8.015	0.155	3.374	0.643	3.360	0.645	3.363	0.644	3.370	0.643
6	10.031	0.123	9.048	0.171	6.228	0.398	5.883	0.437	6.404	0.380	6.138	0.408
7	10.295	0.172	9.141	0.243	6.928	0.437	6.539	0.478	7.023	0.426	6.739	0.457
8	10.300	0.245	9.174	0.328	6.928	0.544	6.551	0.586	7.039	0.533	6.796	0.559
9	10.584	0.305	9.620	0.382	8.728	0.463	8.700	0.465	8.788	0.457	8.910	0.446
10	10.585	0.391	9.630	0.474	8.734	0.558	8.873	0.544	8.795	0.552	8.983	0.534
Durbin-Watson				2.008	Durbin-Watson			2.009	Durbin-Watson			2.005
Pr<DW				0.551	Pr<DW			0.555	Pr<DW			0.533
Pr>DW				0.449	Pr>DW			0.445	Pr>DW			0.467
ID-TD5	log Wnor				log V				log Cxv			
	Durbin-Watson			0.865	Durbin-Watson			1.056	Durbin-Watson			1.182
	Pr<DW			0.000	Pr<DW			0.000	Pr<DW			0.000
	Pr>DW			1.000	Pr>DW			1.000	Pr>DW			1.000
	AR(1) Autocorrelation			0.568	AR(1) Autocorrelation			0.472	AR(1) Autocorrelation			0.409
	Standard error			0.026	Standard error			0.028	Standard error			0.029
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	0.500	0.480	0.486	0.486	0.476	0.491	0.461	0.497	3.422	0.064	3.429	0.064
2	3.668	0.160	3.702	0.157	2.048	0.359	2.048	0.359	4.948	0.084	4.722	0.094
3	3.671	0.299	3.703	0.295	2.048	0.562	2.052	0.562	5.262	0.154	5.237	0.155
4	3.675	0.452	3.703	0.448	2.065	0.724	2.061	0.725	6.016	0.198	6.026	0.197
5	3.677	0.597	3.703	0.593	2.346	0.800	2.311	0.805	6.050	0.301	6.089	0.298
6	3.702	0.717	3.746	0.711	2.354	0.884	2.323	0.888	6.089	0.413	6.106	0.412
7	3.873	0.794	3.866	0.795	2.507	0.927	2.466	0.930	6.474	0.486	6.452	0.488
8	6.968	0.540	7.258	0.509	4.561	0.803	4.676	0.792	6.666	0.573	6.616	0.579
9	6.981	0.639	7.258	0.610	4.565	0.871	4.680	0.861	6.863	0.651	6.945	0.643
10	7.257	0.701	7.306	0.696	5.189	0.878	5.026	0.889	7.556	0.672	7.710	0.657
AR(1) Durbin-Watson				1.934	AR(1) Durbin-Watson			2.007	AR(1) Durbin-Watson			2.024
AR(1) Pr<DW				0.143	AR(1) Pr<DW			0.540	AR(1) Pr<DW			0.642
AR(1) Pr>DW				0.857	AR(1) Pr>DW			0.460	AR(1) Pr>DW			0.358

Table 4.2: DW, Q and LM tests for TD-ID-BIG5 fit to GenTD-BIG data.

ID-TD10	X				Zxv				T			
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	1.006	0.316	1.066	0.302	0.451	0.502	0.439	0.508	0.678	0.410	0.663	0.415
2	5.350	0.069	5.487	0.064	0.501	0.779	0.489	0.783	1.173	0.556	1.169	0.557
3	5.459	0.141	5.532	0.137	0.505	0.918	0.495	0.920	1.175	0.759	1.172	0.760
4	6.567	0.161	6.390	0.172	1.073	0.899	1.073	0.899	1.511	0.825	1.492	0.828
5	7.035	0.218	6.745	0.240	1.827	0.873	1.852	0.869	2.629	0.757	2.632	0.757
6	8.279	0.218	7.638	0.266	3.939	0.685	3.928	0.686	5.113	0.529	5.189	0.520
7	8.488	0.292	7.722	0.358	4.724	0.694	4.740	0.692	6.098	0.528	6.339	0.501
8	8.489	0.387	7.775	0.456	4.779	0.781	4.791	0.780	6.130	0.633	6.402	0.602
9	8.818	0.454	8.234	0.511	6.505	0.689	6.685	0.670	8.040	0.530	8.356	0.499
10	8.833	0.548	8.258	0.604	6.635	0.759	7.016	0.724	8.087	0.620	8.550	0.575
Durbin-Watson				2.012	Durbin-Watson			2.014	Durbin-Watson			2.010
Pr<DW				0.576	Pr<DW			0.588	Pr<DW			0.560
Pr>DW				0.424	Pr>DW			0.412	Pr>DW			0.440
ID-TD10	log Wnor				log V				log Cxv			
	Durbin-Watson			0.901	Durbin-Watson			1.218	Durbin-Watson			1.331
	Pr<DW			0.000	Pr<DW			0.000	Pr<DW			0.000
	Pr>DW			1.000	Pr>DW			1.000	Pr>DW			1.000
	AR(1) Autocorrelation			0.550	AR(1) Autocorrelation			0.391	AR(1) Autocorrelation			0.334
	Standard error			0.027	Standard error			0.029	Standard error			0.030
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	0.154	0.695	0.144	0.704	0.144	0.704	0.134	0.714	0.371	0.543	0.382	0.537
2	1.100	0.577	1.069	0.586	0.239	0.887	0.219	0.896	1.029	0.598	1.049	0.592
3	1.902	0.593	1.882	0.597	1.340	0.720	1.303	0.729	1.340	0.720	1.398	0.706
4	1.913	0.752	1.888	0.756	1.428	0.839	1.387	0.846	2.435	0.656	2.508	0.643
5	2.188	0.823	2.153	0.828	1.447	0.919	1.412	0.923	2.751	0.738	2.811	0.729
6	2.306	0.890	2.306	0.890	1.578	0.954	1.554	0.956	3.713	0.716	3.611	0.729
7	2.317	0.940	2.354	0.938	1.724	0.974	1.769	0.972	3.748	0.808	3.641	0.820
8	7.847	0.449	8.512	0.385	4.343	0.825	4.669	0.792	3.817	0.873	3.734	0.880
9	8.027	0.531	8.568	0.478	4.382	0.885	4.671	0.862	3.845	0.921	3.776	0.926
10	8.202	0.609	8.994	0.533	4.431	0.926	4.796	0.904	3.877	0.953	3.814	0.955
AR(1) Durbin-Watson				1.911	AR(1) Durbin-Watson			2.008	AR(1) Durbin-Watson			2.007
AR(1) Pr<DW				0.077	AR(1) Pr<DW			0.542	AR(1) Pr<DW			0.537
AR(1) Pr>DW				0.923	AR(1) Pr>DW			0.458	AR(1) Pr>DW			0.463

Table 4.3: DW, Q and LM tests for TD-ID-BIG10 fit to GenTD-BIG data.

ID-TD30	X				Zxv				T			
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	0.992	0.319	1.051	0.305	0.020	0.888	0.016	0.899	0.341	0.559	0.333	0.564
2	5.191	0.075	5.324	0.070	0.338	0.845	0.353	0.838	0.823	0.663	0.817	0.665
3	5.289	0.152	5.363	0.147	0.341	0.952	0.354	0.950	1.282	0.733	1.290	0.732
4	6.326	0.176	6.168	0.187	1.088	0.896	1.120	0.891	1.445	0.836	1.430	0.839
5	6.791	0.237	6.528	0.258	1.090	0.955	1.121	0.952	1.518	0.911	1.507	0.912
6	7.957	0.241	7.366	0.288	2.616	0.855	2.502	0.868	2.166	0.904	2.133	0.907
7	8.150	0.320	7.442	0.384	3.253	0.861	3.122	0.874	4.701	0.696	4.760	0.689
8	8.153	0.419	7.507	0.483	3.336	0.912	3.167	0.924	5.178	0.738	5.356	0.719
9	8.497	0.485	7.972	0.537	4.049	0.908	3.915	0.917	5.507	0.788	5.876	0.752
10	8.520	0.578	8.005	0.628	4.067	0.944	3.965	0.949	5.516	0.854	5.876	0.826
Durbin-Watson				2.007	Durbin-Watson			1.999	Durbin-Watson			1.983
Pr<DW				0.543	Pr<DW			0.491	Pr<DW			0.396
Pr>DW				0.457	Pr>DW			0.510	Pr>DW			0.604
ID-TD30	log Wnor				log V				log Cxv			
	Durbin-Watson			0.800	Durbin-Watson			1.506	Durbin-Watson			1.661
	Pr<DW			0.000	Pr<DW			0.000	Pr<DW			0.000
	Pr>DW			1.000	Pr>DW			1.000	Pr>DW			1.000
	AR(1) Autocorrelation			0.601	AR(1) Autocorrelation			0.245	AR(1) Autocorrelation			0.169
	Standard error			0.026	Standard error			0.031	Standard error			0.031
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	1.875	0.171	1.914	0.167	1.931	0.165	1.950	0.163	0.381	0.537	0.384	0.536
2	2.869	0.238	2.993	0.224	2.181	0.336	2.157	0.340	0.453	0.798	0.452	0.798
3	2.999	0.392	3.055	0.383	2.793	0.425	2.803	0.423	0.600	0.897	0.600	0.897
4	3.315	0.507	3.419	0.490	2.852	0.583	2.826	0.587	0.854	0.931	0.828	0.935
5	3.338	0.648	3.447	0.632	3.104	0.684	3.004	0.699	1.282	0.937	1.202	0.945
6	3.656	0.723	3.907	0.689	3.573	0.734	3.377	0.760	1.955	0.924	1.798	0.937
7	3.791	0.804	4.022	0.777	10.836	0.146	11.281	0.127	3.593	0.825	3.557	0.829
8	5.089	0.748	5.700	0.681	10.879	0.209	11.282	0.186	3.595	0.892	3.557	0.895
9	6.477	0.691	7.153	0.621	12.158	0.205	12.733	0.175	3.840	0.922	3.768	0.926
10	6.878	0.737	7.963	0.632	12.647	0.244	13.739	0.185	4.259	0.935	4.190	0.938
AR(1) Durbin-Watson				1.846	AR(1) Durbin-Watson			2.011	AR(1) Durbin-Watson			2.007
AR(1) Pr<DW				0.007	AR(1) Pr<DW			0.567	AR(1) Pr<DW			0.542
AR(1) Pr>DW				0.993	AR(1) Pr>DW			0.433	AR(1) Pr>DW			0.458

Table 4.4: DW, Q and LM tests for TD-ID-BIG30 fit to GenTD-BIG data.

ID-TD65	X				Zxv				T			
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	2.488	0.115	2.593	0.107	1.060	0.303	1.076	0.300	0.035	0.851	0.035	0.851
2	8.716	0.013	8.648	0.013	4.538	0.103	4.423	0.110	0.065	0.968	0.065	0.968
3	9.077	0.028	8.785	0.032	4.655	0.199	4.477	0.214	0.116	0.990	0.117	0.990
4	10.798	0.029	9.991	0.041	6.089	0.193	5.628	0.229	0.117	0.998	0.118	0.998
5	11.856	0.037	10.710	0.057	6.195	0.288	5.665	0.340	0.145	1.000	0.145	1.000
6	13.371	0.038	11.602	0.072	8.402	0.210	7.308	0.293	0.148	1.000	0.147	1.000
7	13.537	0.060	11.615	0.114	8.540	0.287	7.337	0.395	0.178	1.000	0.177	1.000
8	13.550	0.094	11.768	0.162	8.632	0.374	7.667	0.467	0.200	1.000	0.199	1.000
9	13.955	0.124	12.350	0.194	9.632	0.381	8.833	0.453	0.225	1.000	0.225	1.000
10	14.012	0.172	12.418	0.258	10.000	0.441	9.251	0.509	0.255	1.000	0.257	1.000
Durbin-Watson				1.963	Durbin-Watson			1.955	Durbin-Watson			1.972
Pr<DW				0.278	Pr<DW			0.236	Pr<DW			0.329
Pr>DW				0.723	Pr>DW			0.764	Pr>DW			0.671
ID-TD65	log Wnor				log V				log Cxv			
	Durbin-Watson			0.710	Durbin-Watson			1.952	Durbin-Watson			1.980
	Pr<DW			0.000	Pr<DW			0.487	Pr<DW			0.377
	Pr>DW			1.000	Pr>DW			0.224	Pr>DW			0.624
	AR(1) Autocorrelation			0.645	AR(1) Autocorrelation			0.024	AR(1) Autocorrelation			0.010
	Standard error			0.024	Standard error			0.032	Standard error			0.032
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	0.183	0.669	0.193	0.661	0.001	0.979	0.001	0.978	0.000	0.997	0.000	0.996
2	0.185	0.912	0.194	0.907	0.001	1.000	0.001	1.000	0.010	0.995	0.010	0.995
3	0.464	0.927	0.528	0.913	1.164	0.762	1.167	0.761	1.031	0.794	1.034	0.793
4	0.595	0.964	0.622	0.961	2.247	0.690	2.253	0.689	2.617	0.624	2.623	0.623
5	0.727	0.982	0.807	0.977	2.256	0.813	2.260	0.812	2.617	0.759	2.623	0.758
6	1.382	0.967	1.644	0.949	3.242	0.778	3.110	0.795	3.876	0.693	3.728	0.713
7	1.550	0.981	1.729	0.973	3.242	0.862	3.122	0.874	3.893	0.792	3.729	0.811
8	2.898	0.941	3.355	0.910	3.250	0.918	3.123	0.926	3.893	0.867	3.734	0.880
9	2.900	0.968	3.390	0.947	3.255	0.953	3.123	0.959	3.896	0.918	3.735	0.928
10	6.197	0.798	7.457	0.682	4.124	0.942	3.836	0.954	5.112	0.884	4.707	0.910
AR(1) Durbin-Watson				1.807	AR(1) Durbin-Watson			1.998	AR(1) Durbin-Watson			1.999
AR(1) Pr<DW				0.001	AR(1) Pr<DW			0.487	AR(1) Pr<DW			0.496
AR(1) Pr>DW				0.999	AR(1) Pr>DW			0.513	AR(1) Pr>DW			0.504

Table 4.5: DW, Q and LM tests for TD-ID-BIG65 fit to GenTD-BIG data.

4.3.2 IBM Data

We now report on the results of the TD-ID-BIG model fit to the real-world IBM data. We have seen in the previous section that the results of fitting this model to generated 65 minute data are to be considered unreliable. Therefore, we only report on the results of the fits to the 5, 10 and 30 minute data here.

ML estimates

Table 4.6 gives the ML estimates and standard errors obtained from fitting the TD-ID-BIG model to the 5, 10 and 30 minute IBM data. The results of fitting the ID-BIG model of Section 3.2.3 to the IBM data are repeated here for ease of reference. We consider the TD-ID-BIG results first before comparing them to the ID-BIG results. Taking standard errors into account the three sets of TD-ID-BIG estimates agree well, as should be the case. The information gain as more data is used is again clearly visible, since the standard errors resulting from the fit to the 5 minute data are generally smaller than those resulting from the fit to the 10 minute data, which in turn are smaller than those resulting from the fit to the 30 minute data. Note that the estimates of the lag1 autocorrelation are 0.5344, 0.5482 and 0.5082 respectively, which is significantly different from 0, so that the use of the TD-ID-BIG model is justified.

Parameter	ID-BIG30	ID-BIG10	ID-BIG5
β	-0.2175 (0.0601)	-0.2552 (0.0514)	-0.2912 (0.0490)
ψ	1.5821 (0.0786)	1.6871 (0.0698)	1.8118 (0.0728)
ν	0.0035 (0.0010)	0.0042 (0.0008)	0.0050 (0.0007)
ϕ	-0.0346 (0.0290)	-0.0272 (0.0273)	-0.0280 (0.0268)
$\alpha_0 x 10^5$	1.7040 (0.5372)	1.1125 (0.2843)	0.9349 (0.2511)
α_1	0.0964 (0.0149)	0.0960 (0.0102)	0.0950 (0.0088)
β_1	0.8577 (0.0252)	0.8737 (0.0144)	0.8807 (0.0119)
ρ	0.2446 (0.0307)	0.3889 (0.0292)	0.4218 (0.0287)
Parameter	**TD-ID-BIG30**	**TD-ID-BIG10**	**TD-ID-BIG5**
β	-0.1026 (0.0636)	-0.1192 (0.0605)	-0.1643 (0.0603)
ψ	1.5185 (0.0703)	1.6028 (0.0665)	1.7271 (0.0674)
ν	0.0013 (0.0011)	0.0016 (0.0010)	0.0025 (0.0010)
ϕ	-0.0270 (0.0302)	-0.0207 (0.0301)	-0.0202 (0.0304)
$\alpha_0 x 10^5$	0.8803 (0.3618)	0.6787 (0.2800)	0.5713 (0.2604)
α_1	0.0432 (0.0113)	0.0505 (0.0098)	0.0559 (0.0091)
β_1	0.9338 (0.0172)	0.9332 (0.0131)	0.9319 (0.0117)
ρ	0.5082 (0.0525)	0.5482 (0.0379)	0.5344 (0.0354)

Table 4.6: ML estimates of ID-BIG and TD-ID-BIG fits to IBM data.

Comparing the ID-BIG and TD-ID-BIG ML estimates

We now compare the ML estimates obtained from fitting the ID-BIG model to the IBM data, as reported on in Table 3.8 of Section 3.3.3 and repeated in Table 4.6 for ease of reference, to the ML estimates obtained when fitting the TD-ID-BIG model to the same IBM data, as reported on above. Again, as motivated in (A.2.10), for the ID-BIG models we have estimated ρ by the autocorrelation in the estimated $\log \hat{W}_{XV,n}$'s. As in Section 4.3.1, we are interested in the extent to which the estimates obtained from fitting the ID-BIG model to data with dependence in the W_n's will be influenced by this dependence. Taking standard errors into account, it is clear that the estimates of α_1 obtained from the ID-BIG fits are larger than those obtained from the TD-ID-BIG fits, while the estimates of β_1 and ρ obtained from the ID-BIG fits are smaller than those obtained from the TD-ID-BIG fit. Furthermore, this discrepancy becomes substantial when lower frequency data is used. These results are in line with the corresponding results of Section 4.3.1, where it was found that the estimates of α_1, β_1 and ρ are severely biased by the dependence in the W_n's. Note, that the differences between the estimates of α_0 and β obtained by the ID-BIG and TD-ID-BIG models also seem quite substantial, but the standard errors of these estimates are quite large, making it implausible to attribute the differences to the dependence in the W_n's. Hence, as anticipated in Section 4.3.1, the results obtained from the ID-BIG model fit to the IBM data do indeed differ significantly from the results obtained from fitting the more realistic TD-ID-BIG model.

Quality of fit

Figures 4.19 to 4.24 show the PIT and Q-Q plots of the $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ residuals and the $\hat{W}_{NOR,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ estimates. All the plots are reasonably close to the equi-angular line, so that the residuals and estimates correspond well to their model implied distributions. Thus, the TD-ID-BIG model fits the IBM data well.

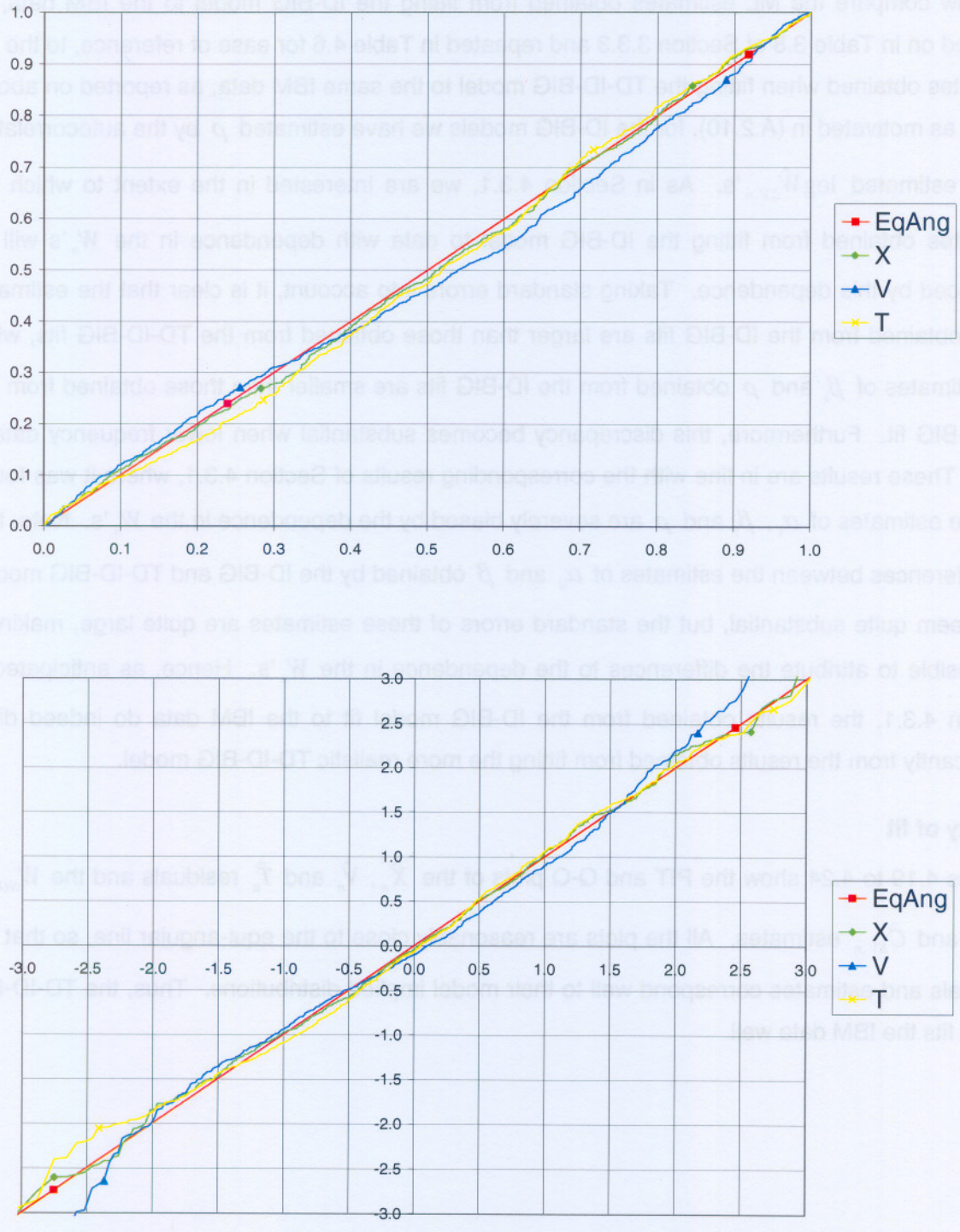

Figure 4.19: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking TD-ID-BIG5 fit to IBM data.

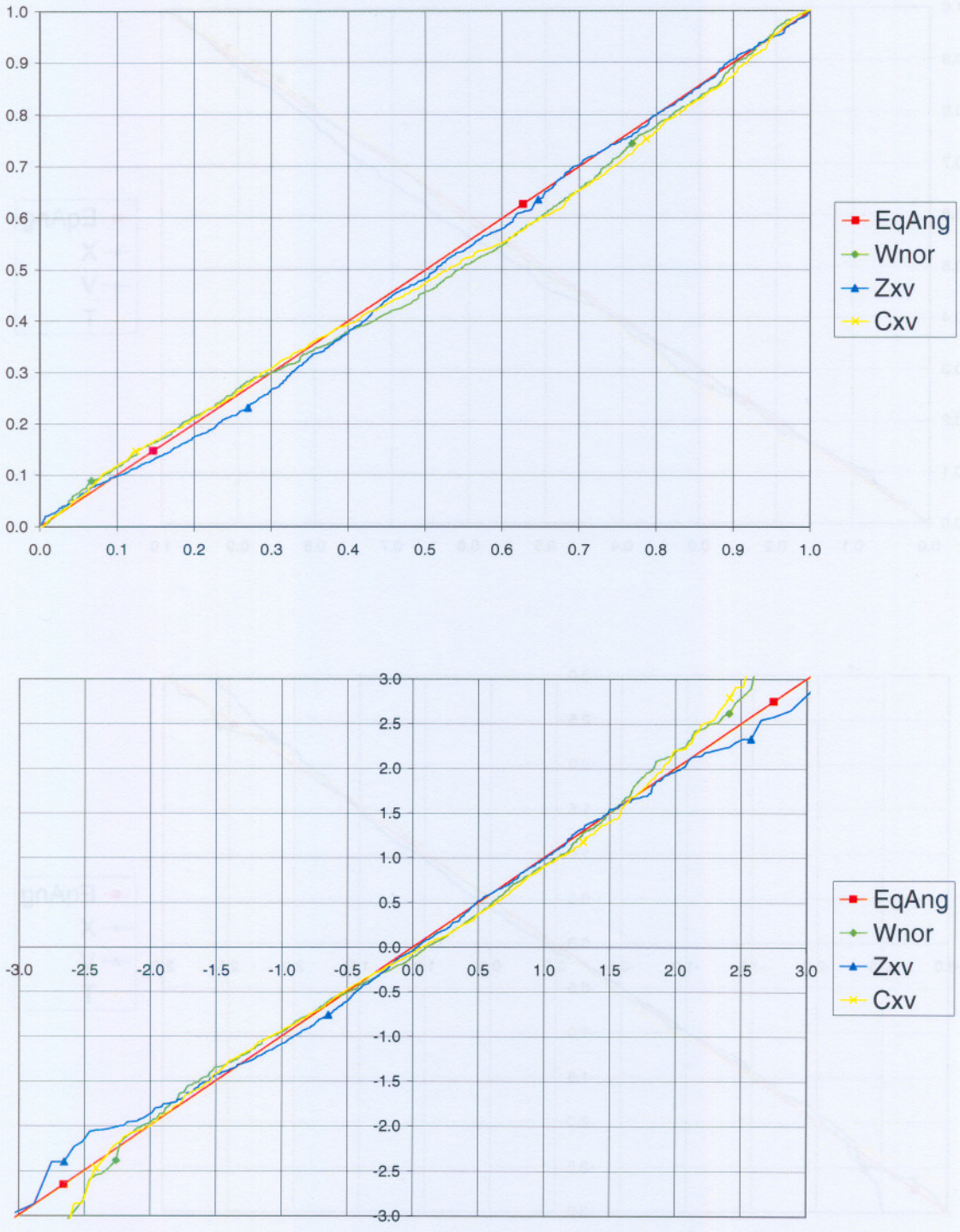

Figure 4.20: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{NOR,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking TD-ID-BIG5 fit to IBM data.

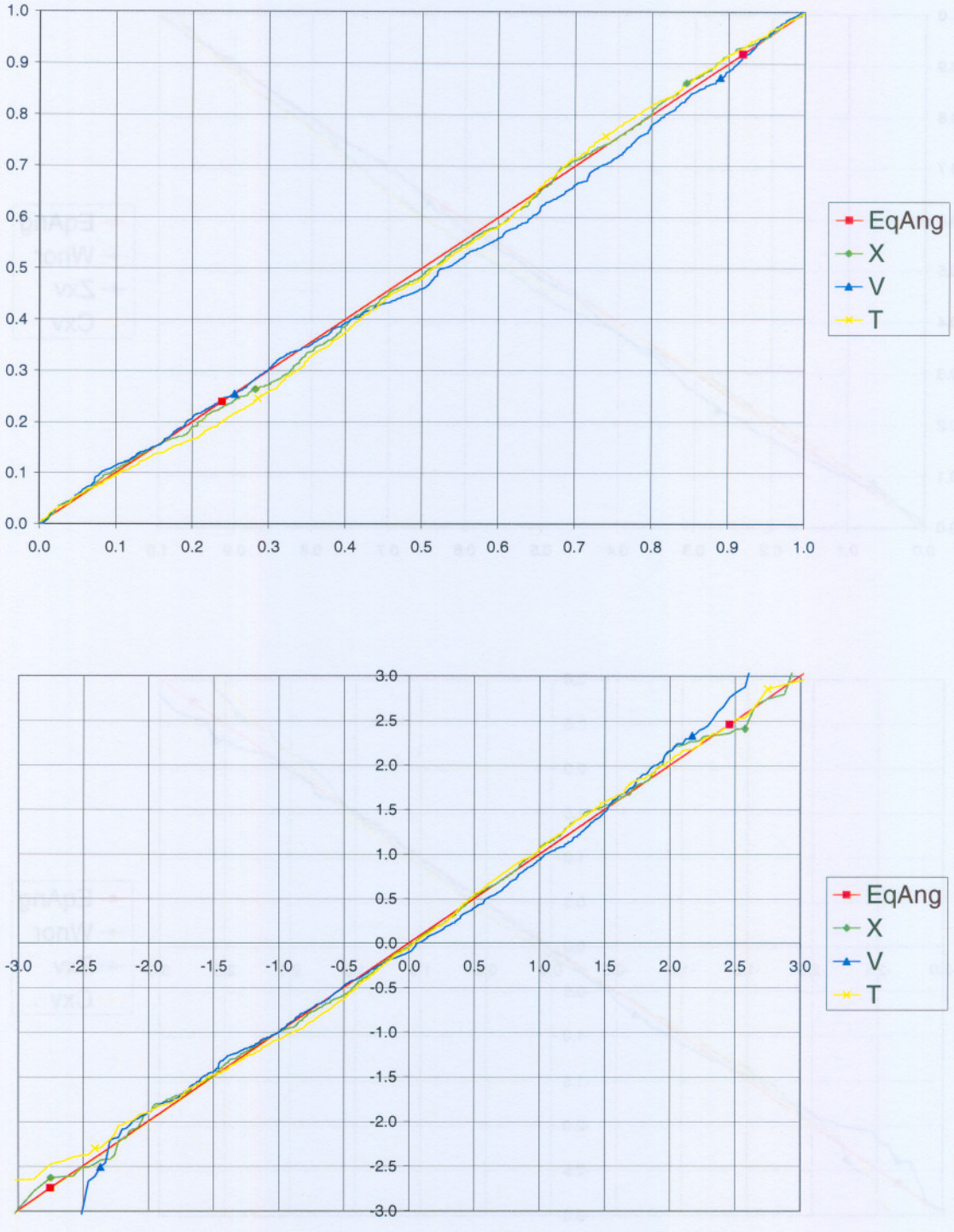

Figure 4.21: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking TD-ID-BIG10 fit to IBM data.

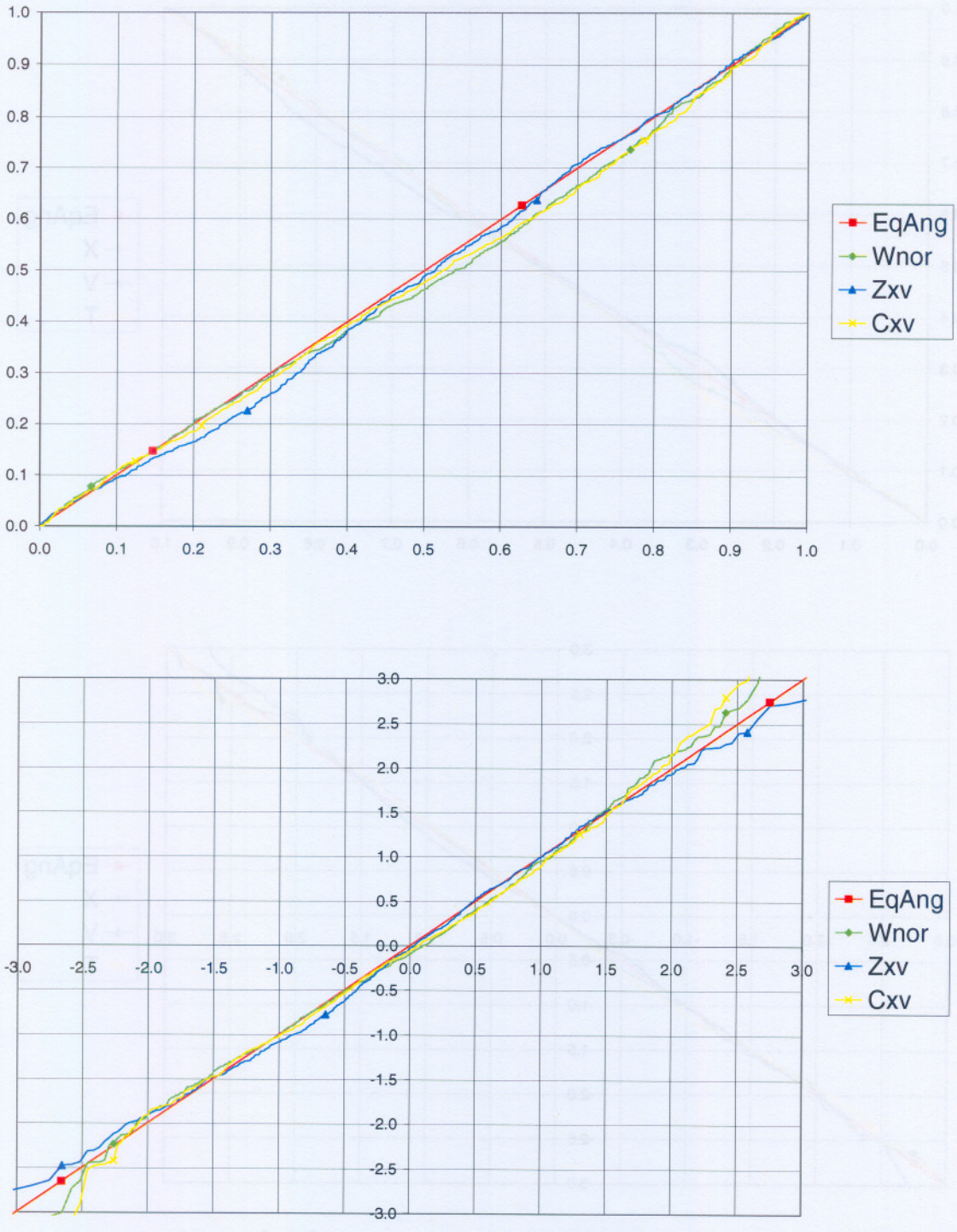

Figure 4.22: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{NOR,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking TD-ID-BIG10 fit to IBM data.

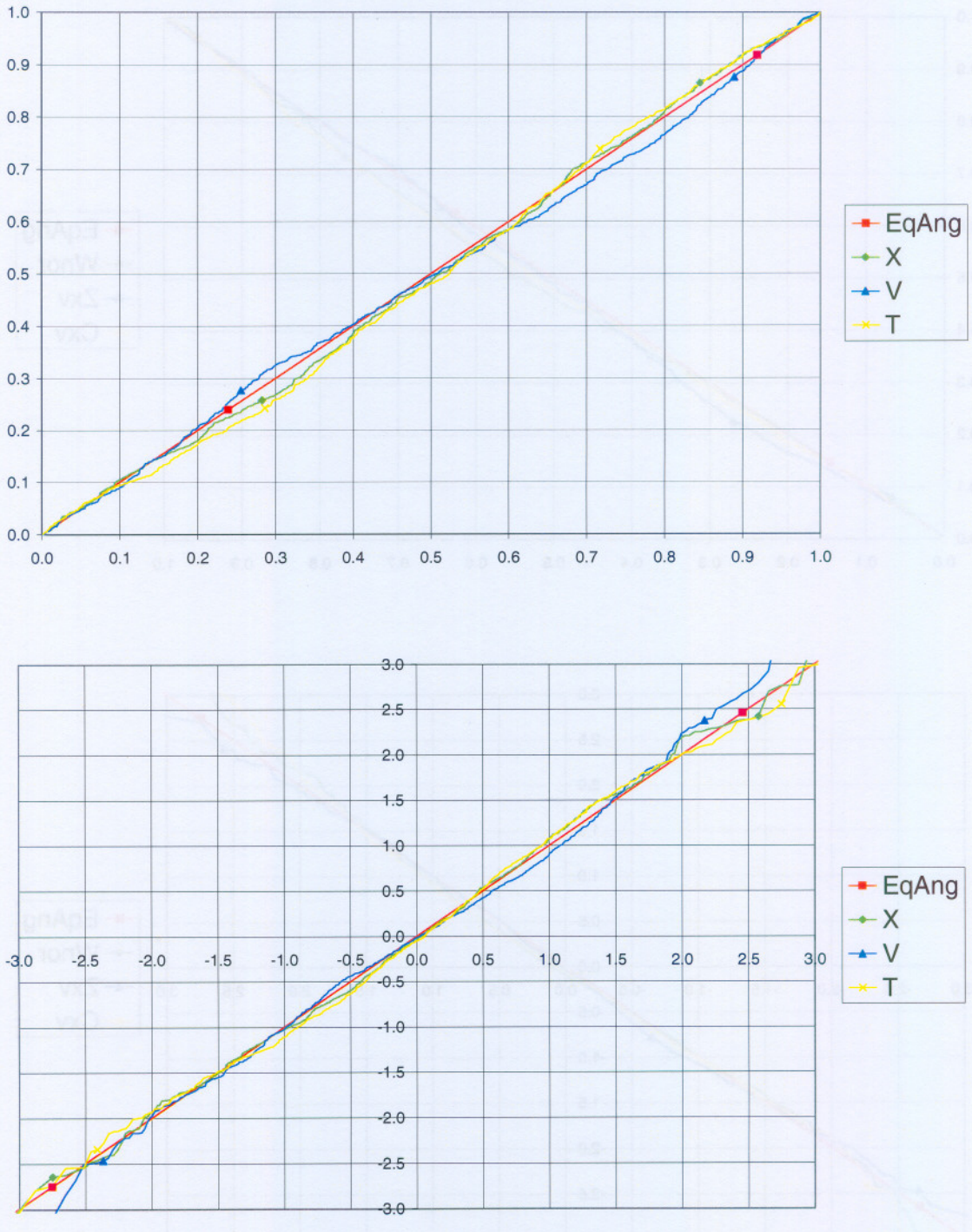

Figure 4.23: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{X}_n$, $\hat{V}_n$ and $\hat{T}_n$ checking TD-ID-BIG30 fit to IBM data.

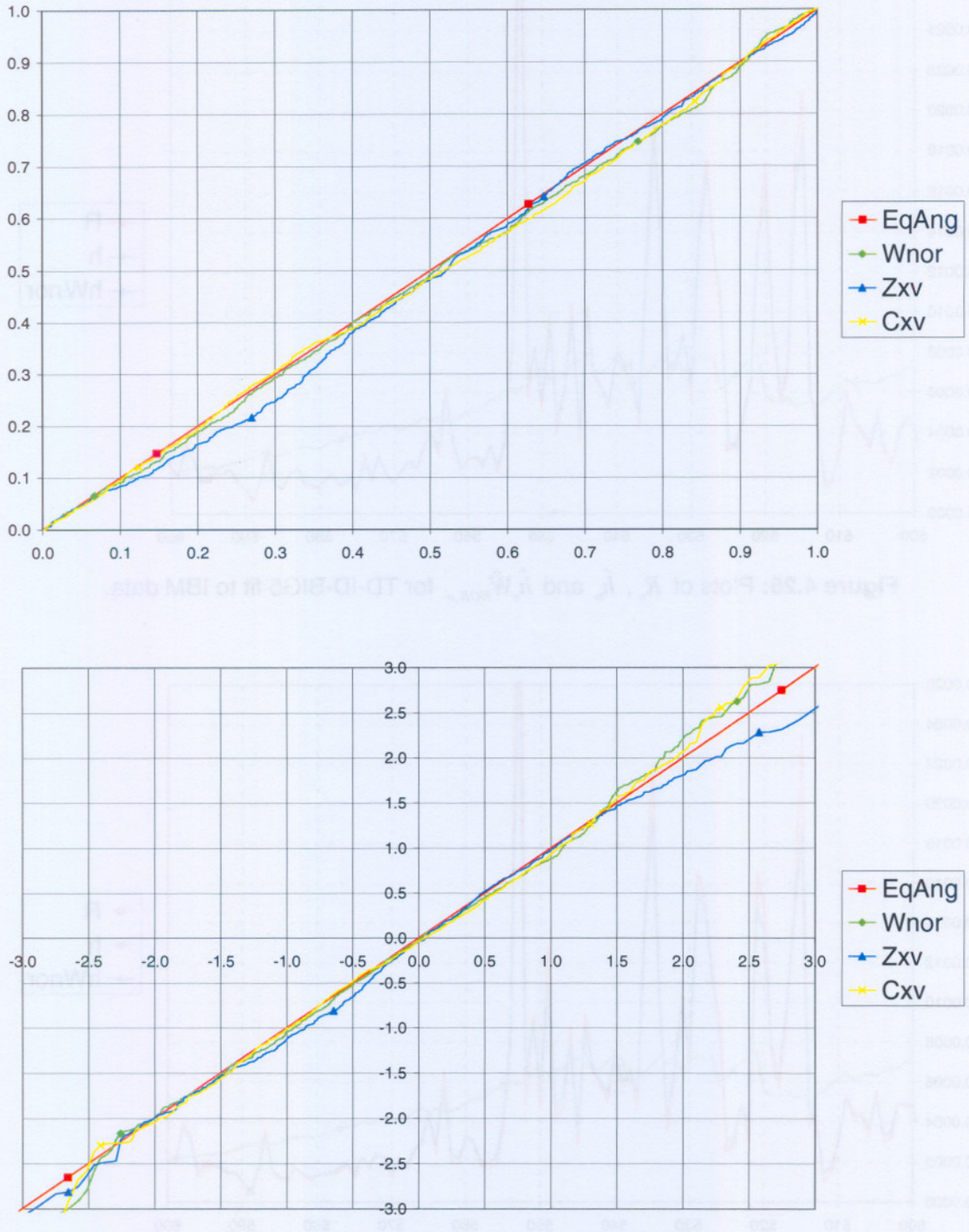

Figure 4.24: PIT (top panel) and Q-Q (bottom panel) plots of $\hat{W}_{NOR,n}$, $\hat{Z}_{XV,n}$ and $\hat{C}_{XV,n}$ checking TD-ID-BIG30 fit to IBM data.

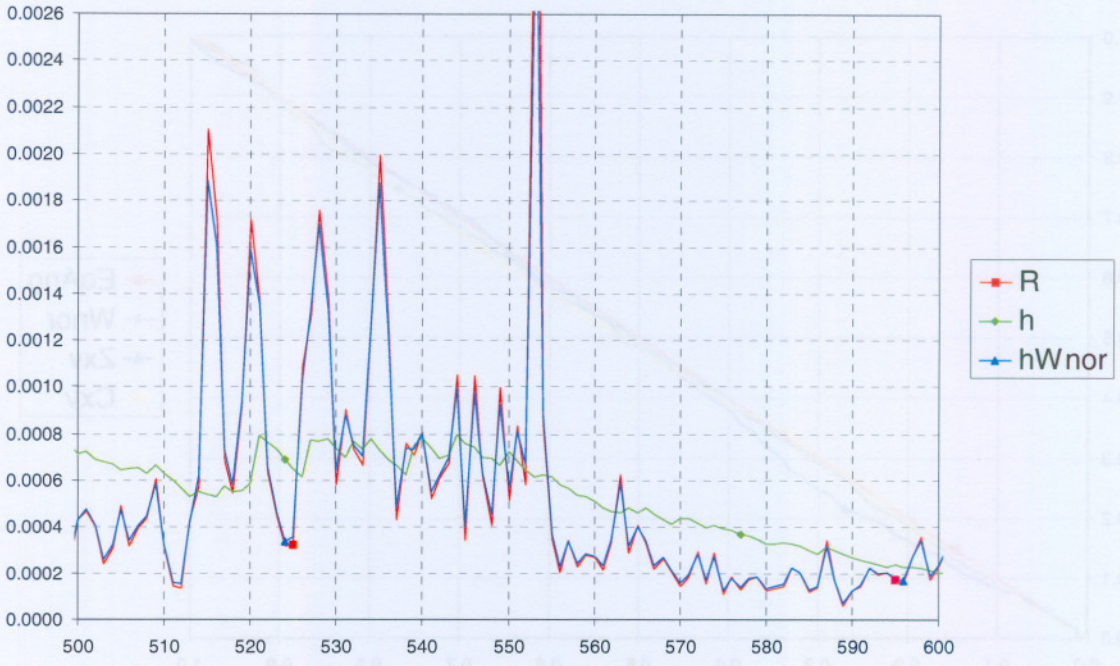

Figure 4.25: Plots of R_n, $\hat{h}_n$ and $\hat{h}_n\hat{W}_{NOR,n}$ for TD-ID-BIG5 fit to IBM data.

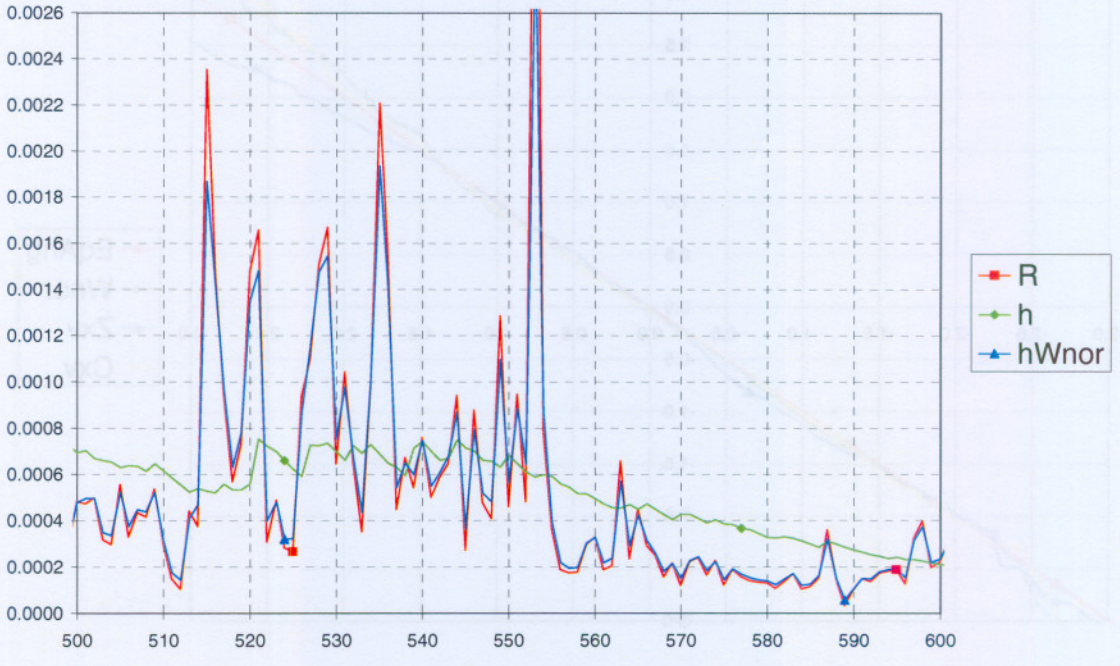

Figure 4.26: Plots of R_n, $\hat{h}_n$ and $\hat{h}_n\hat{W}_{NOR,n}$ for TD-ID-BIG10 fit to IBM data.

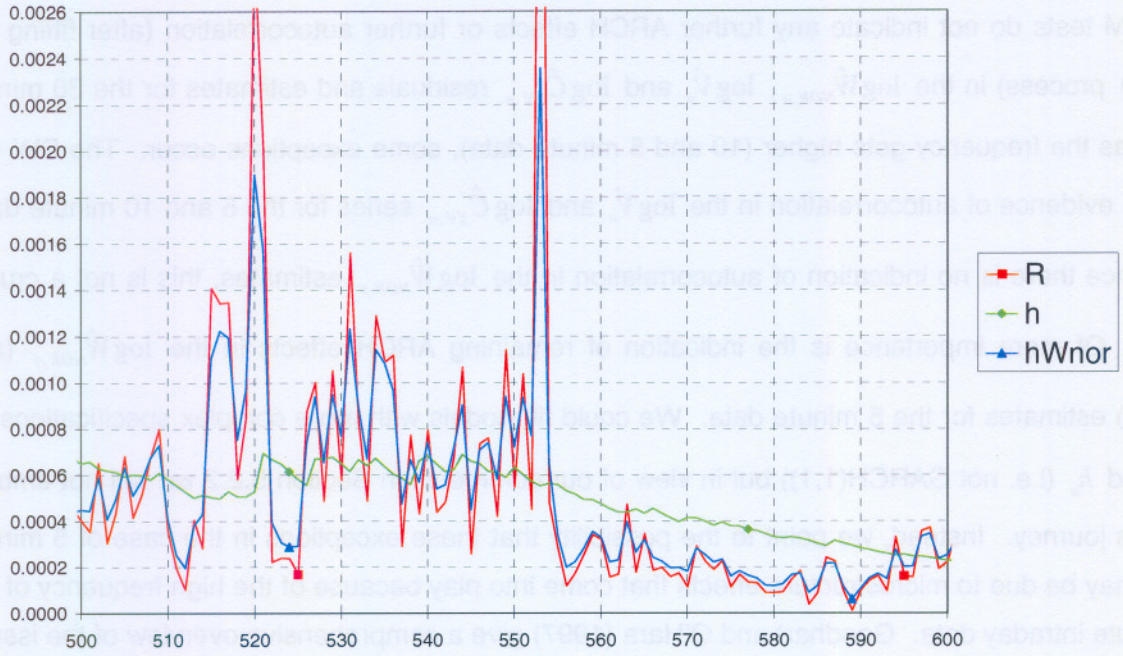

Figure 4.27: Plots of R_n, $\hat{h}_n$ and $\hat{h}_n\hat{W}_{NOR,n}$ for TD-ID-BIG30 fit to IBM data.

Volatility estimates

Figures 4.25 to 4.27 display $\hat{h}_n$, $\hat{h}_n\hat{W}_{NOR,n}$ and R_n for the respective fits. Again, we see that the ordinary GARCH volatility estimate $\hat{h}_n$ is a smoothed estimate of the unknown daily actual volatility, while $\hat{h}_n\hat{W}_{NOR,n}$ is more volatile and tracks the realised volatility R_n much closer than $\hat{h}_n$ does. Figures 4.29 to 4.31 also show the extend to which $\hat{h}_n\hat{W}_{NOR,n}$ adjusts $\hat{h}_n$ for the 5, 10 and 30 minute data respectively. Again, we see that $\hat{h}_n\hat{W}_{NOR,n}$ is closer to R_n when more-frequent intraday data is used, indicating the accuracy gain.

Testing for independence

Finally, we report the results of the DW, Q and LM tests on the residuals and estimates resulting from the TD-ID-BIG fit to the 5, 10 and 30 minute IBM data in Tables 4.7 to 4.9. The DW, Q and LM tests do not indicate any further ARCH effects or autocorrelation in the $\hat{X}_n$ and $\hat{T}_n$ residuals and $\hat{Z}_{XV,n}$ estimates for the 5, 10 and 30 minute data. As expected, the DW test strongly indicates the presence of autocorrelation in the $\log\hat{W}_{NOR,n}$, $\log\hat{V}_n$ and $\log\hat{C}_{XV,n}$ residuals and estimates for

the 5, 10 and 30 minute data, and we therefore fit an $AR(1)$ on these series. While the DW, Q and LM tests do not indicate any further ARCH effects or further autocorrelation (after fitting the $AR(1)$ process) in the $\log \hat{W}_{NOR,n}$, $\log \hat{V}_n$ and $\log \hat{C}_{XV,n}$ residuals and estimates for the 30 minute data, as the frequency gets higher (10 and 5 minute data), some exceptions occur. The DW test shows evidence of autocorrelation in the $\log \hat{V}_n$ and $\log \hat{C}_{XV,n}$ series for the 5 and 10 minute data, but since there is no indication of autocorrelation in the $\log \hat{W}_{NOR,n}$ estimates, this is not a crucial point. Of more importance is the indication of remaining ARCH effects in the $\log \hat{W}_{NOR,n}$ (and $\log \hat{V}_n$) estimates for the 5 minute data. We could fit models with more complex specifications for μ_n and h_n (i.e. not GARCH(1,1)) but in view of our comments in Section 5.2.2 we will not embark on this journey. Instead, we point to the possibility that these exceptions in the case of 5 minute data may be due to microstructure effects that come into play because of the high frequency of the 5 minute intraday data. Goodhart and O'Hara (1997) give a comprehensive overview of the issues surrounding the use of high frequency data in financial applications. We set aside the study of these microstructure effects in our context as future work.

4.4 Conclusion

In this chapter we introduced the TD-ID-BIG model, an extension the ID-BIG model of Section 3.2.3, to allow for dependence in the random impact factors $\{W_n\}$. Since an explicit equation for the likelihood function of this TD-ID-BIG model is unavailable, we obtained an approximation by using EIS. We also applied the control variate technique to reduce the variance of the likelihood function approximation.

We fit the TD-ID-BIG model to generated and empirical data sets and compared the results to the results of fitting the ID-BIG model to these data sets. We found that the TD-ID-BIG model fits the data well as long as enough intraday data (with sufficiently high frequency) is available. Some of the parameter estimates obtained by fitting the ID-BIG model to the data sets are, however, biased in the presence of such dependence and these results should not be regarded as adequate. Evidently it is necessary to entertain the somewhat more complicated TD-ID-BIG model which allows for dependence.

We conclude this chapter with a final remark about the random impact factors $\{W_n\}$. Recall that we think of W_n as representing the impact of news noise on volatility over and above volatility

expected from events on prior days. We would expect the W_n's obtained by fitting the TD-ID-BIG model to two different stock return series to be positively correlated if the price of the stocks in question are expected to be similarly influenced by such news noise. For instance, given the nature of their business, we would expect the share prices of Wal-Mart and Home Depot (both constituents of the Dow Jones index) to react fairly similarly to news that is not company specific. To see whether this expectation is correct we fitted the TD-ID-BIG model to these two share return series, each comprising of 5 minute returns for 927 consecutive trading days from 16 February 2000 to 24 October 2003. Figure 4.28 shows the scatter plot of the two resultant series of $\left\{\log \hat{W}_{NOR,n}\right\}$ estimates. The positive correlation of 0.5806 between the two series can be clearly seen. This agrees with our expectation to a large extent and seems to support our interpretation of the role of the W_n's.

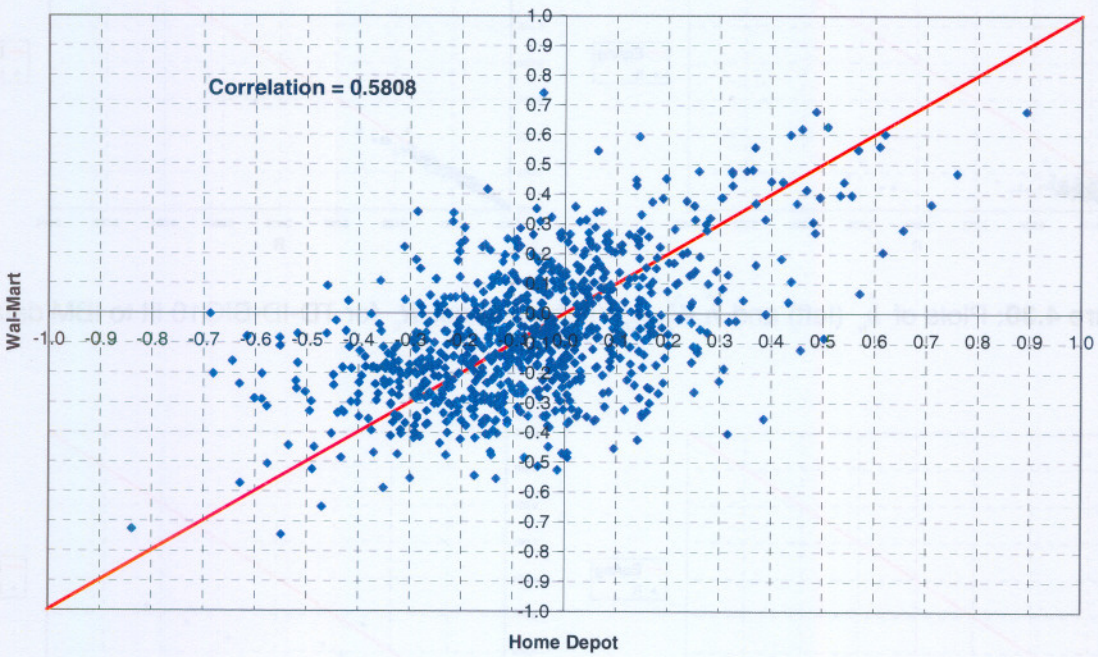

Figure 4.28: Scatter plot of $\log \hat{W}_{NOR,n}$ for TD-ID-BIG5 fit to Wal-Mart and Home Depot data.

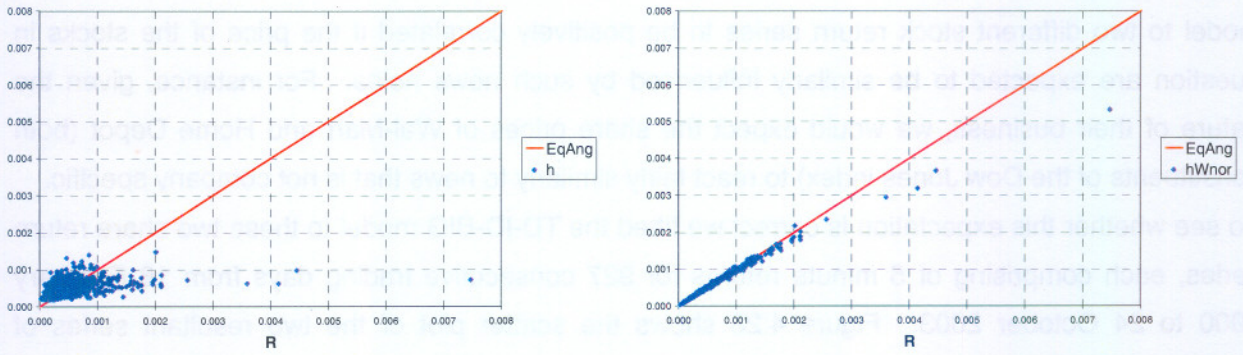

Figure 4.29: Plots of $\hat{h}_n$ (left) and $\hat{h}_n\hat{W}_{NOR,n}$ (right) against R_n for TD-ID-BIG5 fit to IBM data.

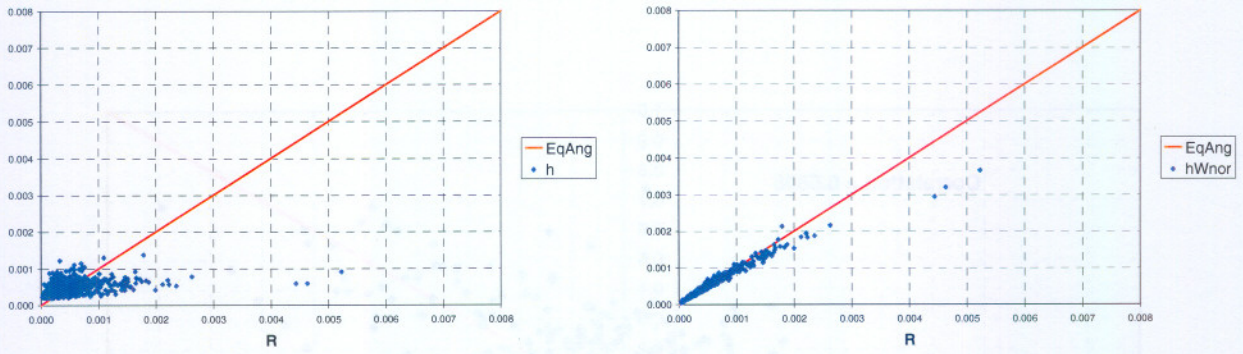

Figure 4.30: Plots of $\hat{h}_n$ (left) and $\hat{h}_n\hat{W}_{NOR,n}$ (right) against R_n for TD-ID-BIG10 fit to IBM data.

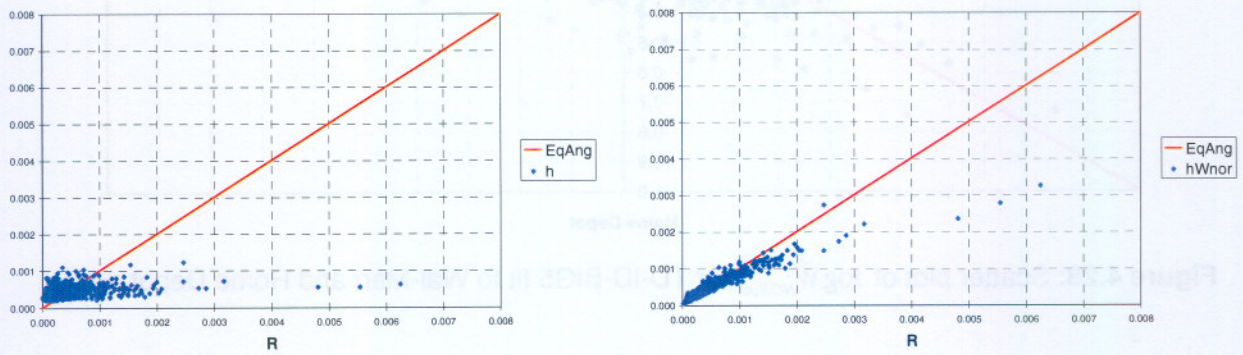

Figure 4.31: Plots of $\hat{h}_n$ (left) and $\hat{h}_n\hat{W}_{NOR,n}$ (right) against R_n for TD-ID-BIG30 fit to IBM data.

TD-ID-BIG5	X				Zxv				T			
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	0.232	0.630	0.237	0.626	1.094	0.296	1.109	0.292	0.647	0.421	0.658	0.417
2	0.279	0.870	0.272	0.873	1.094	0.579	1.114	0.573	0.698	0.706	0.683	0.711
3	2.424	0.489	2.565	0.464	1.941	0.585	2.234	0.525	1.300	0.729	1.493	0.684
4	2.440	0.655	2.613	0.625	2.313	0.678	2.868	0.580	1.475	0.831	1.829	0.767
5	2.883	0.718	3.017	0.697	3.122	0.681	3.408	0.637	2.137	0.830	2.299	0.806
6	3.038	0.804	3.241	0.778	3.341	0.765	3.558	0.736	2.506	0.868	2.582	0.859
7	4.432	0.729	4.495	0.721	3.348	0.851	3.559	0.829	2.520	0.926	2.583	0.921
8	4.507	0.809	4.620	0.797	3.565	0.894	3.723	0.881	2.747	0.949	2.748	0.949
9	5.270	0.810	5.503	0.789	3.586	0.937	3.730	0.928	2.762	0.973	2.750	0.973
10	5.322	0.869	5.535	0.853	4.208	0.938	4.575	0.918	3.410	0.970	3.636	0.962
Durbin-Watson				2.013	Durbin-Watson			1.966	Durbin-Watson			1.962
Pr<DW				0.584	Pr<DW			0.295	Pr<DW			0.273
Pr>DW				0.416	Pr>DW			0.705	Pr>DW			0.727
TD-ID-BIG5	log Wnor				log V				log Cxv			
	Durbin-Watson			0.841	Durbin-Watson			1.034	Durbin-Watson			1.216
	Pr<DW			0.000	Pr<DW			0.000	Pr<DW			0.000
	Pr>DW			1.000	Pr>DW			1.000	Pr>DW			1.000
	AR(1) Autocorrelation			0.580	AR(1) Autocorrelation			0.483	AR(1) Autocorrelation			0.392
	Standard error			0.026	Standard error			0.028	Standard error			0.029
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	0.141	0.708	0.148	0.701	0.084	0.772	0.079	0.779	1.959	0.162	1.963	0.161
2	10.345	0.006	10.393	0.006	7.998	0.018	8.033	0.018	3.267	0.195	3.135	0.209
3	10.986	0.012	11.122	0.011	8.783	0.032	8.705	0.034	3.295	0.348	3.203	0.361
4	11.221	0.024	11.164	0.025	8.992	0.061	8.750	0.068	3.349	0.501	3.265	0.515
5	11.440	0.043	11.242	0.047	9.238	0.100	8.848	0.115	3.352	0.646	3.265	0.659
6	11.750	0.068	11.604	0.071	9.589	0.143	9.266	0.159	3.867	0.695	3.738	0.712
7	11.908	0.104	11.683	0.112	9.780	0.201	9.389	0.226	4.482	0.723	4.275	0.748
8	12.151	0.145	12.061	0.149	10.038	0.262	9.750	0.283	4.483	0.811	4.293	0.830
9	13.219	0.153	13.249	0.152	10.708	0.296	10.541	0.309	4.483	0.877	4.294	0.891
10	13.633	0.190	13.859	0.180	11.358	0.330	11.350	0.331	4.701	0.910	4.551	0.919
AR(1) Durbin-Watson				2.083	Durbin-Watson			2.123	Durbin-Watson			2.124
AR(1) Pr<DW				0.901	Pr<DW			0.973	Pr<DW			0.974
AR(1) Pr>DW				0.099	Pr>DW			0.027	Pr>DW			0.026

Table 4.7: DW, Q and LM tests for TD-ID-BIG5 fit to IBM data.

TD-ID-BIG10	X				Zxv				T			
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	0.126	0.723	0.130	0.718	0.956	0.328	0.963	0.327	0.470	0.493	0.473	0.492
2	0.218	0.897	0.205	0.903	1.338	0.512	1.459	0.482	0.722	0.697	0.795	0.672
3	2.027	0.567	2.155	0.541	3.650	0.302	4.316	0.229	2.598	0.458	3.081	0.379
4	2.036	0.729	2.186	0.702	4.189	0.381	5.255	0.262	2.851	0.583	3.549	0.471
5	2.563	0.767	2.679	0.749	4.916	0.426	5.626	0.344	3.417	0.636	3.870	0.568
6	2.676	0.848	2.849	0.828	5.738	0.453	6.160	0.406	4.780	0.572	4.902	0.556
7	4.357	0.738	4.376	0.736	5.763	0.568	6.224	0.514	4.804	0.684	4.954	0.666
8	4.457	0.814	4.526	0.807	6.242	0.620	6.685	0.571	5.374	0.717	5.517	0.701
9	5.142	0.822	5.337	0.804	6.447	0.695	6.875	0.650	5.710	0.769	5.839	0.756
10	5.174	0.879	5.354	0.866	7.112	0.715	7.735	0.655	6.315	0.788	6.612	0.762
Durbin-Watson				2.020	AR(1) Durbin-Watson			1.979	AR(1) Durbin-Watson			1.968
Pr<DW				0.621	AR(1) Pr<DW			0.372	AR(1) Pr<DW			0.305
Pr>DW				0.379	AR(1) Pr>DW			0.628	AR(1) Pr>DW			0.695
TD-ID-BIG10	log Wnor				log V				log Cxv			
	Durbin-Watson			0.763	Durbin-Watson			1.093	Durbin-Watson			1.279
	Pr<DW			0.000	Pr<DW			0.000	Pr<DW			0.000
	Pr>DW			1.000	Pr>DW			1.000	Pr>DW			1.000
	AR(1) Autocorrelation			0.618	AR(1) Autocorrelation			0.453	AR(1) Autocorrelation			0.360
	Standard error			0.025	Standard error			0.028	Standard error			0.030
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	0.385	0.535	0.394	0.530	0.011	0.917	0.009	0.923	0.192	0.661	0.193	0.660
2	1.160	0.560	1.178	0.555	0.177	0.916	0.188	0.910	0.374	0.829	0.375	0.829
3	1.238	0.744	1.258	0.739	0.319	0.957	0.309	0.958	0.451	0.929	0.440	0.932
4	2.622	0.623	2.661	0.616	1.550	0.818	1.566	0.815	1.202	0.878	1.206	0.877
5	3.194	0.670	3.227	0.665	2.125	0.832	2.071	0.839	1.368	0.928	1.387	0.926
6	5.605	0.469	5.621	0.467	3.719	0.715	3.655	0.723	1.709	0.944	1.669	0.948
7	5.762	0.568	5.674	0.578	3.902	0.791	3.779	0.805	2.073	0.956	1.988	0.961
8	5.762	0.674	5.685	0.683	4.045	0.853	3.939	0.863	2.120	0.977	2.031	0.980
9	9.215	0.418	9.463	0.396	6.335	0.706	6.453	0.694	2.165	0.989	2.078	0.990
10	9.314	0.503	9.528	0.483	6.773	0.747	6.769	0.747	2.375	0.993	2.265	0.994
AR(1) Durbin-Watson				2.088	AR(1) Durbin-Watson			2.146	AR(1) Durbin-Watson			2.123
AR(1) Pr<DW				0.914	AR(1) Pr<DW			0.989	AR(1) Pr<DW			0.974
AR(1) Pr>DW				0.086	AR(1) Pr>DW			0.011	AR(1) Pr>DW			0.026

Table 4.8: DW, Q and LM tests for TD-ID-BIG10 fit to IBM data.

TD-ID-BIG30	X				Zxv				T			
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	0.038	0.846	0.041	0.840	0.303	0.582	0.327	0.568	0.733	0.392	0.710	0.400
2	0.169	0.919	0.152	0.927	1.096	0.578	1.259	0.533	1.702	0.427	1.804	0.406
3	1.519	0.678	1.633	0.652	2.857	0.414	3.441	0.329	2.812	0.422	2.925	0.403
4	1.531	0.821	1.664	0.797	3.742	0.442	4.705	0.319	3.173	0.529	3.317	0.506
5	2.184	0.823	2.281	0.809	4.124	0.532	4.864	0.433	3.637	0.603	3.682	0.596
6	2.230	0.897	2.360	0.884	4.172	0.654	4.865	0.561	3.850	0.697	3.784	0.706
7	4.473	0.724	4.418	0.731	4.202	0.756	4.868	0.676	3.851	0.797	3.785	0.804
8	4.620	0.797	4.610	0.798	4.587	0.801	5.197	0.736	3.950	0.862	3.901	0.866
9	5.177	0.819	5.283	0.809	4.826	0.849	5.463	0.792	3.976	0.913	3.933	0.916
10	5.179	0.879	5.283	0.872	4.893	0.898	5.505	0.855	3.986	0.948	3.944	0.950
Durbin-Watson				2.017	AR(1) Durbin-Watson			1.999	AR(1) Durbin-Watson			1.981
Pr<DW				0.608	AR(1) Pr<DW			0.492	AR(1) Pr<DW			0.382
Pr>DW				0.392	AR(1) Pr>DW			0.508	AR(1) Pr>DW			0.618
TD-ID-BIG30	log Wnor				log V				log Cxv			
	Durbin-Watson			0.715	Durbin-Watson			1.395	Durbin-Watson			1.561
	Pr<DW			0.000	Pr<DW			0.000	Pr<DW			0.000
	Pr>DW			1.000	Pr>DW			1.000	Pr>DW			1.000
	AR(1) Autocorrelation			0.642	AR(1) Autocorrelation			0.303	AR(1) Autocorrelation			0.220
	Standard error			0.024	Standard error			0.030	Standard error			0.031
Order	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM	Q	Pr>Q	LM	Pr>LM
1	0.019	0.890	0.017	0.897	1.545	0.214	1.514	0.219	0.196	0.658	0.190	0.663
2	0.329	0.849	0.342	0.843	1.694	0.429	1.640	0.440	0.462	0.794	0.449	0.799
3	0.468	0.926	0.458	0.928	1.694	0.638	1.642	0.650	0.705	0.872	0.693	0.875
4	0.752	0.945	0.720	0.949	5.990	0.200	5.771	0.217	3.396	0.494	3.280	0.512
5	4.527	0.476	4.272	0.511	6.082	0.298	5.810	0.325	8.322	0.139	8.192	0.146
6	6.721	0.347	6.338	0.386	8.412	0.210	7.780	0.255	10.323	0.112	10.143	0.119
7	6.955	0.434	6.674	0.464	9.015	0.252	8.486	0.292	12.826	0.077	12.243	0.093
8	7.805	0.453	7.642	0.469	9.440	0.307	8.795	0.360	12.857	0.117	12.377	0.135
9	8.327	0.502	8.115	0.523	9.924	0.357	9.542	0.389	12.869	0.169	12.397	0.192
10	8.522	0.578	8.409	0.589	10.319	0.413	10.017	0.439	13.155	0.215	12.960	0.226
AR(1) Durbin-Watson				2.018	AR(1) Durbin-Watson			2.093	AR(1) Durbin-Watson			2.055
AR(1) Pr<DW				0.603	AR(1) Pr<DW			0.928	AR(1) Pr<DW			0.806
AR(1) Pr>DW				0.398	AR(1) Pr>DW			0.072	AR(1) Pr>DW			0.194

Table 4.9: DW, Q and LM tests for TD-ID-BIG30 fit to IBM data.

CHAPTER 5

CONCLUDING COMMENTS

5.1 Looking Back

Chapter 1 introduced the main concepts this thesis deals with, namely volatility (and realised volatility); GARCH models; the use of alternative innovation distributions to account for the heavier tails often encountered in financial returns series; and the use of high, low and close data, as well as full intraday data, instead of only daily data in GARCH models in order to obtain more accurate parameter estimates. Chapter 2 dealt with probabilistic aspects regarding IG-related distributions in preparation for implementation in Chapters 3 and 4. In Chapter 3 we extended the basic daily normal GARCH model by assuming NIG-distributed innovations (or BIG processes) instead of normally distributed innovations (or BM processes). We also catered for the use of high, low and close prices, as well as full intraday data, instead of only daily data. We found that the models allowing for innovations with heavier tails and catering for full intraday data fit empirical data sets better than models assuming normally distributed innovations and making use of only daily data. However, we also found disturbing indications of time dependence in the random impact factors, which we assumed to be independent. This finding casts doubt on the validity of the results in the presence of such time dependence. In Chapter 4 we extended the ID-BIG model and introduced the TD-ID-BIG model, which allows for dependence in the random impact factors. We obtained an approximation for the likelihood function of this TD-ID-BIG model by using EIS, since the traditional GARCH likelihood expression is no longer available in this case. We also applied the control variate technique to reduce the variance of this likelihood function approximation. Fitting this new time dependent model to generated and empirical data sets, we found that the TD-ID-BIG model fits the data well as long as enough intraday data (with sufficiently high frequency) is available. We also found that some of the ID-BIG parameter estimates are biased in the presence of such time dependence, making the TD-ID-BIG extension mandatory.

Finally, we draw attention to the fact that some of the research presented in Chapters 2 and 3 has already been published as a paper (see Venter, de Jongh and Griebenow (2005)), while another paper (see Venter de Jongh and Griebenow (2004)) is currently under consideration for publication.

5.2 Looking Forward

Research work for a thesis must stop (or at least pause) somewhere in order to complete the thesis document. This has the unfortunate consequence that one is aware of many open issues raised by the research in progress – issues that require further thinking but have to be left dangling as it were, until further work can be done. We close by listing some of these topics.

5.2.1 Intraday Random Impact Factors

In the models presented in Chapters 3 and 4, we used $UIG(\psi)$-distributed random impact factors W_n to model factors that may have an effect on the volatility experienced on day n, such as news arriving after close of market on day $n-1$. These news events may occur overnight, before opening of the market on day n, but they often arise during the day as opposed to in-between market closing and market opening. When the market opens on day n, the market players would have had maybe several hours to digest an overnight event. On the other hand, a news report arriving three hours after market opening would not have had an effect on the first three hours' trading, but will have an effect on the trading during the rest of the day (and probably the following day). Thus, it is to be expected that news arriving overnight will influence trading differently from news arriving during the trading day. This suggests dividing every day into, say, two periods. The random impact factors representing the overnight news events are then supposed to influence trading mainly during the first period of the day (with its influence diminishing thereafter), while the random impact factors representing the news events occurring during the day are supposed to influence trading mainly during the second period of the day.

We extended the ID-BIG model of Section 3.2.3 to include two random impact factor periods per day and carried out standard maximum likelihood inference in order to obtain ML estimates and corresponding standard errors for the parameters. Preliminary empirical studies have shown that, taking standard errors into account, the estimates of ψ are indeed significantly different for the two random impact factor periods. This indicates that the effects of news events occurring overnight are indeed different than the effects of news events occurring during the trading day. This is presented as a topic for further research and we do not report any further results in this thesis.

5.2.2 Other Specifications of μ_n and h_n

A myriad of other choices for the structural (μ_n) and volatility (h_n) parts of the GARCH specification may be examined. Numerous such alternative specifications abound in the literature

(see the surveys mentioned in Chapter 1). However, in his Nobel lecture, Engle (2004) states the following:

"The GARCH(1,1) specification is the workhorse of financial applications. It is remarkable that one model can be used to describe the volatility dynamics of almost any financial return series."

This statement, supported by the results given in Hansen and Lunde (2001), where it is found that the GARCH(1,1) specification is a very competitive specification, leads us to expect no drastic improvements in our results when using more complicated GARCH specifications. Nevertheless, we would not really know if this is indeed the case until further work has been done.

5.2.3 More Complex Processes for W_n

In Chapter 4 we assumed the dependence in the random impact factors W_n to be described by an $AR(1)$ process, which turned out to be sufficient in our example (IBM return series). However, in other cases a more complex process may be called for to describe this dependence. For this purpose, the TD-ID-BIG model of Section 4.2 could be extended by changing the specification of W_n and revising the derivations to take account of these changes. Similar EIS and variance reduction techniques can then be used to obtain an approximation to the likelihood function and estimate the ML parameters.

5.2.4 Trading Volumes

Finally, incorporating trading volumes into our framework could prove to be a very fruitful exercise. Liesenfeld (2001) analyses the joint dynamic behaviour of stock price volatility and trading volume as being simultaneously affected by the number of information (or news) arrivals and traders' sensitivity to new information. Both these latent processes are specified as serially correlated random variables, which means that high-dimensional integration is required to calculate the likelihood function. This situation is similar to our situation in Chapter 4, and Liesenfeld (2001) also uses a SML approach in order to estimate the model parameters. He finds that both the number of news arrivals and the traders' sensitivity to new information are important factors which influence the price volatility, while sensitivity to new information is largely irrelevant for the behaviour of trading volume. Furthermore, the news arrival process largely directs the short-term volatility movement while the sensitivity to news largely directs the long-term volatility movement. This work relates more to stochastic volatility models than GARCH models and as such were outside the initial scope of our research. However, at this stage of our work it would be interesting to attempt closer alignments.

Fleming, Kirby and Ostdiek (2004) uses the mixture-of-distributions-hypothesis (MDH) to examine the relation between stochastic volatility, volume dynamics and ARCH effects in stock returns. They specify a bivariate stochastic autoregressive volatility (SARV) representation for daily returns and volume in which volume captures a component of volatility that is unrelated to ARCH effects and find indications of strong positive correlation between nonpersistent components of squared returns and trading volume. This suggests that trade volumes could be used to improve on return volatility estimates obtained by using conventional GARCH models. Xu, Chen and Wu (2005) examine the effect of market activity, as measured by the duration between two consecutive trades, on volatility and volume dynamics. They find that shorter duration between trades (or higher trading intensity) implies higher return volatility. Again, it is tempting to dig deeper into these papers and to develop possible relations with our work or extensions of our work. For the time being it must be left for future research.

APPENDIX

A.1 Q-Q plots and PIT plots

Throughout Sections 3.3 and 4.3 we assess the goodness of fit of several models by using Quantile-Quantile (or Q-Q) plots and Probability Integral Transform (or PIT) plots, which we briefly review here for ease of reference.

Q-Q plots are well known and can be used for several purposes. One of these purposes is to test the quality of fit of a theoretical distribution on a data set. The quantiles of the data set are plotted against the quantiles of the theoretical distribution. If the two sets of quantiles agree perfectly, the Q-Q plot would lie entirely on the equi-angular line through the origin. Deviations from the equi-angular line indicate differences in the two sets of quantiles and thus differences between the data set and theoretical distribution.

Another way to determine whether a data set comes from a theoretical distribution is by means of PIT plots. Suppose we want to determine whether the observations $x_1,\ldots,x_N$ come from a distribution with distribution function $F(x)$. We order the observations $x_{(1)} \le \ldots \le x_{(N)}$ and set $U_{(n)} = F\left(x_{(n)}\right)$, $n=1\ldots,N$. If the observations $x_1,\ldots,x_N$ indeed come from a distribution with distribution function $F(x)$, then $U_{(n)}$, $n=1,\ldots,N$ will be distributed like the order statistics of a uniformly distributed random variable on the unit interval and $E\left(U_{(n)}\right) = n/(N+1)$. This means that the pairs $\left(n/(N+1);U_{(n)}\right)$ should be close to the equi-angular line. Deviations from the equi-angular line again indicate differences between the distribution of the observations $x_1,\ldots,x_N$ and the distribution with distribution function $F(x)$. We shall indicate the equi-angular line by the term "EqAng" in both the Q-Q plots and PIT plots.

When the distribution function $F(x)$ is not analytically available we may be able to estimate it empirically as accurately as required (as shown in Sections 2.3.5, 2.5.3 and 2.7.5 for our cases) by generating a number of observations $X_{(1)} \le \ldots \le X_{(M)}$ from this distribution. Setting $T_0 = 0$ we can then calculate $T_n = \max_{T_{n-1} \le m \le M}\left(m : x_{(m)} \le X_{(n)}\right)$ and $U_{(n)} = T_n/M$ for $n=1,\ldots,N$, and plot $\left(n/(N+1);U_{(n)}\right)$ as before. A disadvantage of the Q-Q and PIT plot based assessment of the quality of fit is that it tends to be rather informal in the sense that no numerical measure of fit is

computed to which a p-value can be attached. However, the use of these plots are widespread in the relevant literature (see e.g. Forsberg (2002)) and we shall follow this tradition.

A.2 Sequences of Dependent Random Variables with UIG Marginals

In Section 4.2 we introduced sequences of dependent random variables $\{W_n\}$ with identical $UIG(\psi)$ marginal distributions, based on $N(0,1)$-distributed variables $\{U_n\}$ with lag1 autocorrelation ρ. We finished up at two unknown functions for the lag1 autocorrelation in the W_n's and the lag1 autocorrelation in the $\log W_n$'s, indicated by ρ_W and $\rho_{\log W}$ respectively. We now set out to approximate these functions and propose an algorithm for generating observations from the $AR(\rho)-UIG(\psi)$ process.

Consider the following log-normal (LN) process related to the U_n's, namely

$$V_n = \exp\left(\tau U_n - \frac{1}{2}\tau^2\right),\ \tau > 0\ . \tag{A.2.1}$$

In general, if $\ln X \sim N(\mu,\sigma^2)$, then X is lognormally distributed with $E(X^k) = \exp(k\mu + k^2\sigma^2/2)$. Therefore, since $\tau U_n - \frac{1}{2}\tau^2 = \ln V_n \sim N\left(-\frac{1}{2}\tau^2, \tau^2\right)$ from (4.2.2) and (A.2.1), V_n is lognormally distributed with $E(V_n^k) = \exp\left(-\frac{k\tau^2}{2} + \frac{k^2\tau^2}{2}\right) = \exp\left(\frac{k\tau^2}{2}(k-1)\right)$, so that

$$\begin{aligned} E(V_n) &= \exp(0) = 1 \\ Var(V_n) &= E(V_n^2) - (E(V_n))^2 = \exp(\tau^2) - 1\ . \end{aligned} \tag{A.2.2}$$

Furthermore, it follows from (4.2.1) and (A.2.4) that $\ln V_n V_{n-1} \sim N\left(-\tau^2, 2\tau^2(1+\rho)\right)$, since $V_n V_{n-1} = \exp\left(\tau\left(\rho U_{n-1} + \sqrt{1-\rho^2}Z_n\right) - \frac{1}{2}\tau^2\right)\exp\left(\tau U_{n-1} - \frac{1}{2}\tau^2\right) = \exp\left(\tau(1+\rho)U_{n-1} + \tau\sqrt{1-\rho^2}Z_n - \tau^2\right)$, so that $E\left((V_n V_{n-1})^k\right) = \exp\left(-k\tau^2 + k^2\tau^2(1+\rho)\right)$ and

$$\begin{aligned} Cov(V_n, V_{n-1}) &= E(V_n V_{n-1}) - E(V_n)E(V_{n-1}) = \exp\left(-\tau^2 + \tau^2(1+\rho)\right) - 1 = \exp(\rho\tau^2) - 1 \\ Corr(V_n, V_{n-1}) &= \frac{Cov(V_n, V_{n-1})}{\sqrt{Var(V_n)}\sqrt{Var(V_{n-1})}} = \frac{\exp(\rho\tau^2) - 1}{\exp(\tau^2) - 1}\ . \end{aligned} \tag{A.2.3}$$

From (2.2.5) and (A.2.5) it follows that, if we select τ such that

$$\exp(\tau^2)-1=1/\psi^2 \text{, i.e. } \tau=\sqrt{\ln(1/\psi^2+1)} \text{ and } \psi=\sqrt{1/(\exp(\tau^2)-1)}, \qquad \text{(A.2.4)}$$

then we have $E(V_n)=E(W_n)$ as well as $Var(V_n)=Var(W_n)$. Since $E(V_n)=1$ we say V_n is unit lognormally (ULN) distributed, and the distribution function of V_n is given by

$$F_{ULN(\tau)}(v)=P(V_n\le v)=P(\exp(\tau U_n-\tau^2/2)\le v)=P\left(U_n\le\frac{\tau}{2}+\frac{\ln v}{\tau}\right)=\Phi\left(\frac{\tau}{2}+\frac{\ln v}{\tau}\right). \qquad \text{(A.2.5)}$$

When (A.2.4) holds, the distribution functions $F_{UIG(\psi)}(w)$ and $F_{ULN(\tau)}(v)$ have the same first two moments and it turns out that they are in fact very similar for $\psi\ge 0.5$, and thus $\tau\le 1.27$, as Figures A.1 and A.2 show. The cases $\psi\ge 1$ ($\tau\le 0.83$) are of most practical interest.

This means that $F_{UIG(\psi)}(v)\approx F_{ULN(\tau)}(v)=\Phi\left(\frac{\tau}{2}+\frac{\ln v}{\tau}\right)$, from (A.2.5). Setting $y=\Phi\left(\frac{\tau}{2}+\frac{\ln v}{\tau}\right)$, we get $v=\exp\left(\tau\Phi^{-1}(y)-\frac{1}{2}\tau^2\right)$, so that $F^{-1}_{UIG(\psi)}(y)\approx\exp\left(\tau\Phi^{-1}(y)-\frac{1}{2}\tau^2\right)$ and therefore

$$W_n=F^{-1}_{UIG(\psi)}(\Phi(U_n))\approx\exp\left(\tau\Phi(\Phi^{-1}(U_n))-\frac{1}{2}\tau^2\right)=\exp\left(\tau U_n-\frac{1}{2}\tau^2\right)=V_n. \qquad \text{(A.2.6)}$$

The joint distribution of $V_1,\ldots,V_N$ may be called the $AR(\rho)-ULN(\tau)$ distribution. This is close to the $AR(\rho)-UIG(\psi)$ distribution, but the latter has the advantage that the marginals are exactly $UIG(\psi)$ so that it fits into our overall NIG-BIG framework. From (A.2.3) and (A.2.4), it seems reasonable that we should have approximately

$$\rho_W=Corr(W_n,W_{n-1})\approx Corr(V_n,V_{n-1})=\frac{\exp(\rho\tau^2)-1}{\exp(\tau^2)-1}=\psi^2\left(\left(1+\frac{1}{\psi^2}\right)^{\rho}-1\right)=\tilde{\rho}. \qquad \text{(A.2.7)}$$

By the approximation $W_n\approx V_n$ we have

$$\log W_n\approx\log V_n=\tau U_n-\frac{1}{2}\tau^2. \qquad \text{(A.2.8)}$$

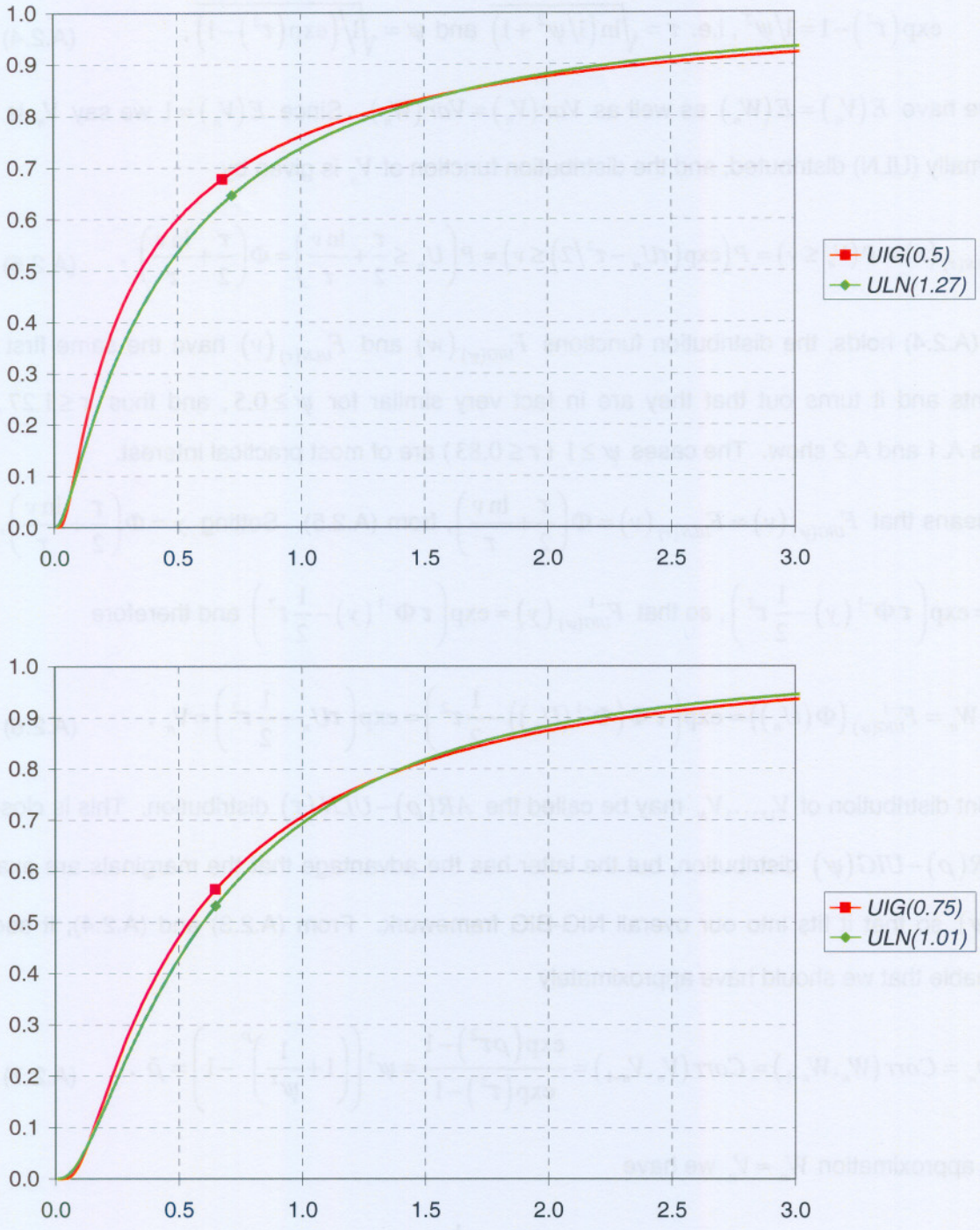

Figure A.1: Comparing $ULN(\tau)$ and $UIG(\psi)$ distributions for $\psi = 0.5$ and $\psi = 0.75$.

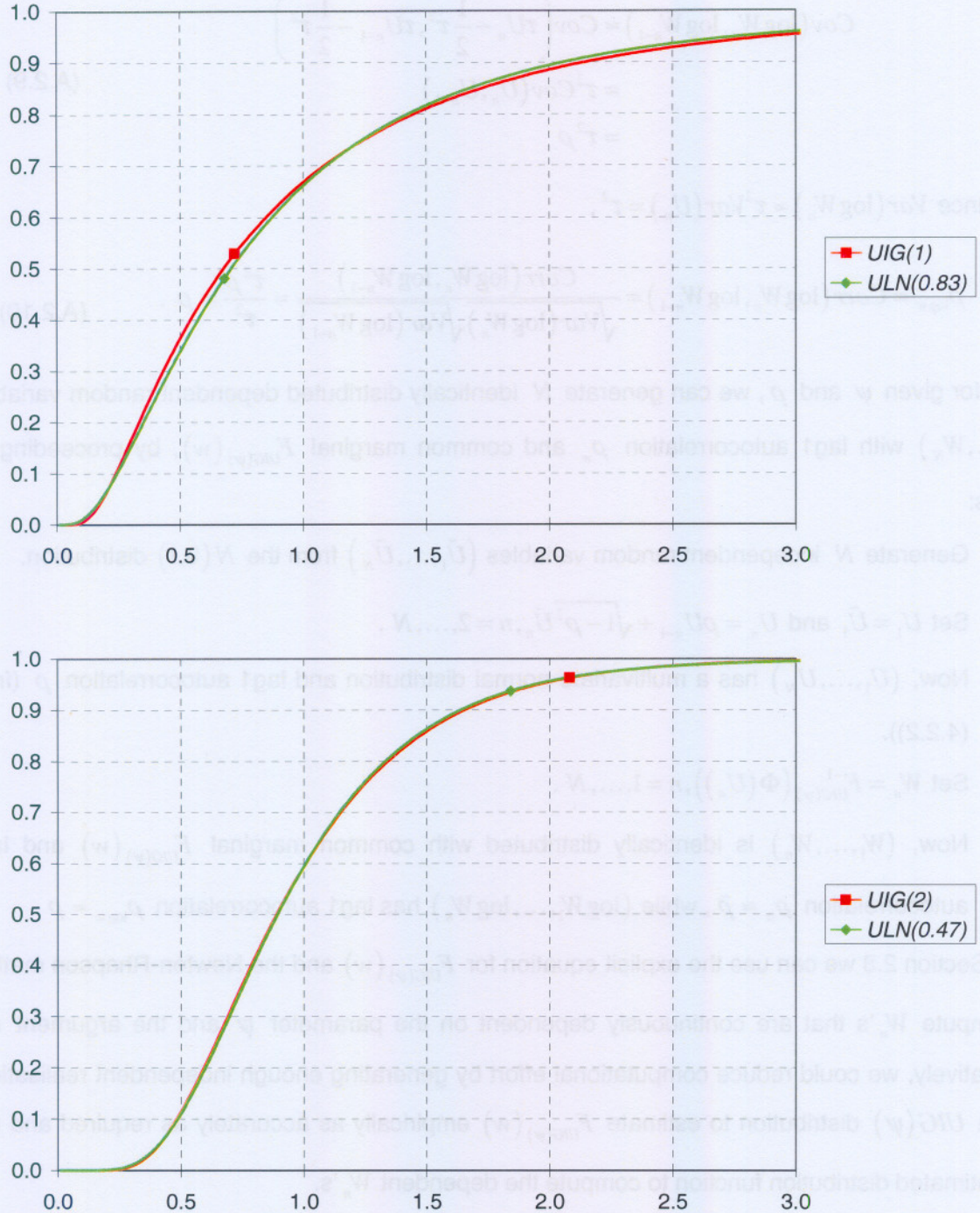

Figure A.2: Comparing $ULN(\tau)$ and $UIG(\psi)$ distributions for $\psi = 1$ and $\psi = 2$.

Hence

$$\begin{aligned} Cov(\log W_n, \log W_{n-1}) &\approx Cov\left(\tau U_n - \frac{1}{2}\tau^2, \tau U_{n-1} - \frac{1}{2}\tau^2\right) \\ &= \tau^2 Cov(U_n, U_{n-1}) \\ &= \tau^2 \rho \end{aligned} \tag{A.2.9}$$

and, since $Var(\log W_n) \approx \tau^2 Var(U_n) = \tau^2$,

$$\rho_{\log w} = Corr(\log W_n, \log W_{n-1}) = \frac{Corr(\log W_n, \log W_{n-1})}{\sqrt{Var(\log W_n)}\sqrt{Var(\log W_{n-1})}} \approx \frac{\tau^2 \rho}{\tau^2} = \rho \ . \tag{A.2.10}$$

Thus, for given ψ and ρ, we can generate N identically distributed dependent random variables $(W_1, \ldots, W_N)$ with lag1 autocorrelation ρ_w and common marginal $F_{UIG(\psi)}(w)$, by proceeding as follows:

- Generate N independent random variables $(\tilde{U}_1, \ldots, \tilde{U}_N)$ from the $N(0,1)$ distribution.
- Set $U_1 = \tilde{U}_1$ and $U_n = \rho U_{n-1} + \sqrt{1-\rho^2}\tilde{U}_n, n = 2, \ldots, N$.
- Now, $(U_1, \ldots, U_N)$ has a multivariate normal distribution and lag1 autocorrelation ρ (from (4.2.2)).
- Set $W_n = F^{-1}_{UIG(\psi)}(\Phi(U_n)), n = 1, \ldots, N$.
- Now, $(W_1, \ldots, W_n)$ is identically distributed with common marginal $F_{UIG(\psi)}(w)$ and lag1 autocorrelation $\rho_w \approx \tilde{\rho}$, while $(\log W_1, \ldots, \log W_n)$ has lag1 autocorrelation $\rho_{\log w} \approx \rho$.

As in Section 2.8 we can use the explicit equation for $F_{UIG(\psi)}(w)$ and the Newton-Rhapson method to compute W_n's that are continuously dependent on the parameter ψ and the argument U_n. Alternatively, we could reduce computational effort by generating enough independent realisations from a $UIG(\psi)$ distribution to estimate $F_{UIG(\psi)}(w)$ empirically as accurately as required and use this estimated distribution function to compute the dependent W_n's.

We will now ascertain how accurately ρ_w and $\rho_{\log w}$ approximate $\tilde{\rho}$ and ρ by means of Monte Carlo simulation. Using half a million independent observations to estimate $F_{UIG(\psi)}(w)$, we generated $N = 20000$ dependent $UIG(\psi)$ observations $(W_1, \ldots, W_{20000})$ for all the combinations of

$\psi = 0.5(0.05)5$ and $\rho = -0.5(0.05)0.5$, as described above, and calculated the empirical lag1 autocorrelation $\hat{\rho}_w$ of the generated W_n's and the empirical lag1 autocorrelation $\hat{\rho}_{\log w}$ of the $\log W_n$'s, for every (ψ, ρ) combination. Figure A.3 shows that $\hat{\rho}_w$ is quite close to $\tilde{\rho}$ (calculated by (A.2.7)) except when $\psi < 1$ and $\rho > 0.2$ simultaneously. Figure A.4 shows that $\hat{\rho}_{\log w}$ estimates ρ very accurately for all the (ψ, ρ) combinations. We therefore choose to focus on the autocorrelation of the $\log W_n$'s rather than the autocorrelation of the W_n's.

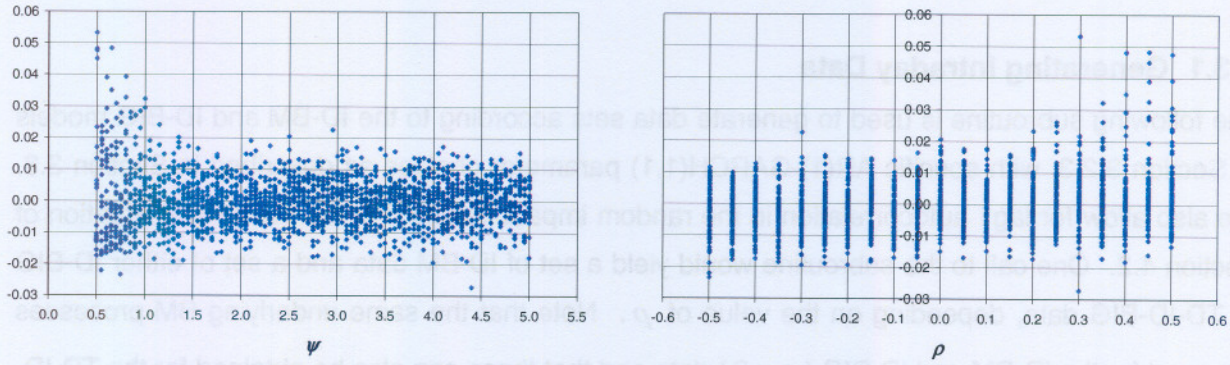

Figure A.3: Scatter plots of $\hat{\rho}_w - \tilde{\rho}$ against ψ (left) and ρ (right).

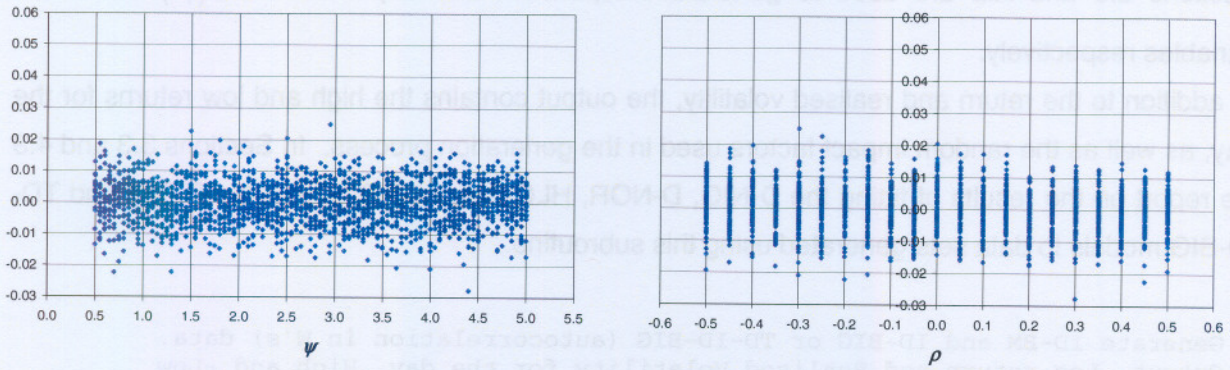

Figure A.4: Scatter plots of $\hat{\rho}_{\log w} - \rho$ against ψ (left) and ρ (right).

A.3 Programs

We will now shed some light on selected parts of the programs we wrote during the completion of this thesis. The programs were written mainly in FORTRAN®, with some of the programs, or parts thereof, replicated in SAS® for verification purposes. We will include only FORTRAN® code in this appendix. As an introduction, Section A.3.1 starts with a subroutine for generating the data that we used to verify our models and to illustrate the underlying theory, as reported on in Sections 3.3 and 4.3. This is followed by a somewhat more complicated subroutine in Section A.3.2, used to calculate the log-likelihood function of the ID-BIG model of Section 3.2.3. Finally, in Section A.3.3 we take an in-depth look at the program for fitting the TD-ID-BIG model of Section 4.2.

A.3.1 Generating Intraday Data

The following subroutine is used to generate data sets according to the ID-BM and ID-BIG models of Section 3.2.3, with specific AR(1)-GARCH(1,1) parameter choices as described in Section 3.3. We also allow for lag1 autocorrelation in the random impact factors, i.e. the TD-ID-BIG situation of Section 4.2. One call to the subroutine would yield a set of ID-BM data and a set of either ID-BIG or TD-ID-BIG data, depending on the value of ρ. Note that the same underlying BM processes are used for the ID-BM and ID-BIG ($\rho = 0$) data and that these can also be obtained for the TD-ID-BIG ($\rho \neq 0$) data by specifying the same random seed. Note that the algorithm described in Sections 2.8 and A.2 are used to generate independent and dependent $UIG(\psi)$-distributed variables respectively.

In addition to the return and realised volatility, the output contains the high and low returns for the day, as well as the random impact factors used in the generation process. In Sections 3.3 and 4.3 we report on the results of fitting the D-NIG, D-NOR, HLC-BM, HLC-BIG, ID-BM, ID-BIG and TD-ID-BIG models to data sets generated using this subroutine.

```
! Generate ID-BM and ID-BIG or TD-ID-BIG (autocorrelation in W's) data.
! Output: Log-return and Realised Volatility for the day, High and -Low
! Log-return and Random Impact Factor (W).
!=======================================================================
      SUBROUTINE GEN_BM_and_BIG_or_TD(tt,ar1,arch0,arch1,garch1,mean,bta,ps,
     & rho,intrv,nintrv,iseed,BIGfile,BMfile)
!=======================================================================
! Declarations
      IMPLICIT NONE
      INTEGER(4) n,t,tt,intrv,nintrv,iseed,steps
      REAL(8) ar1,arch0,arch1,garch1,mean,ps,bta,rho,stepsr,ylag1,ylag2,h,
     &        hlag1,mu,mulag1,rv(5000),lnrho,yhi,ylo,po(5000),pc(5000),
     &        ph(5000),pl(5000),nyhi,nylo,npo(5000),npc(5000),nph(5000),
     &        npl(5000),nyhi,nylo,npo(5000),npc(5000),nph(5000),npl(5000),
```

```
     &        nylag1,nylag2,nhlag1,nh,nmu,nmulag1,nrv(5000)
      REAL(8),DIMENSION(:),ALLOCATABLE::w,b,y,ny
      CHARACTER*(*) BIGfile,BMfile

! Allocate arrays
      IF (n.GT.5000) n=5000
      steps=intrv*nintrv ! 5-min: 78 intervals x 26 steps/interval=2028
      ALLOCATE(w(tt),b(steps),y(steps),ny(steps))
! Generate UIG-distributed random impact factors
      CALL RNSET(iseed)
      IF (rho.EQ.0.0D0) THEN
       CALL GEN_UIG(tt,w,ps,iseed) ! Independent UIGs
      ELSE
       CALL GEN_UIGdep(tt,w,ps,rho,iseed) ! Dependent UIGs
      ENDIF
      iseed=iseed+9999 ! Ensure the same random seed whether rho=0 or not
      CALL RNSET(iseed)
      stepsr=steps
! Starting Lags
      hlag1=arch0/(1.0D0-arch1-garch1); nhlag1=hlag1
      ylag1=mean; ylag2=mean; nylag1=mean; nylag2=mean

! Day-Loop; Start
      DO n=1,tt
! AR(1)-GARCH(1,1)
       IF (n.EQ.1) THEN
        mu=(1.0D0+ar1)*mean
        h=arch0*(1.0D0-arch1)/(1.0D0-arch1-garch1)
        nmu=mu
        nh=h
       ELSE
        mu=mean+ar1*ylag1
        nmu=mean+ar1*nylag1
        h=arch0+arch1*(ylag1-mulag1)**2.0D0+garch1*hlag1
        nh=arch0+arch1*(nylag1-nmulag1)**2.0D0+garch1*nhlag1
       ENDIF
! Generate Brownian Motion
       CALL DRNNOR(steps,b)
       DO t=2,steps; b(t)=b(t-1)+b(t); ENDDO
       DO t=1,steps; b(t)=b(t)/dsqrt(stepsr); ENDDO
! Initiate variables
       yhi=-1.0D-100; ylo=1.0D100; rv(n)=0.0D0
       nyhi=-1.0D-100; nylo=1.0D100; nrv(n)=0.0D0
! Simulate activity on day n; Start
       DO t=1,steps
! Return
        y(t)=mu*t/stepsr+dsqrt(h)*(bta*t*w(n)/stepsr+dsqrt(w(n))*b(t))
        ny(t)=nmu*t/stepsr+dsqrt(nh)*(bta*t/stepsr+b(t))
! Realised Volatility
        IF ((MOD(t,nintrv).EQ.0).AND.(t.GT.nintrv)) THEN
         rv(n)=rv(n)+(y(t)-y(t-nintrv))**2.0D0
         nrv(n)=nrv(n)+(ny(t)-ny(t-nintrv))**2.0D0
        ENDIF
! Max and Min
        IF (y(t).GT.yhi) yhi=y(t)
        IF (y(t).LT.ylo) ylo=y(t)
        IF (ny(t).GT.nyhi) nyhi=ny(t)
        IF (ny(t).LT.nylo) nylo=ny(t)
```

```
      ENDDO
! Simulate activity on day n; End
! Update Lags
      hlag1=h; ylag2=ylag1; ylag1=y(steps); mulag1=mu
      nhlag1=nh;nylag2=nylag1;nylag1=ny(steps);nmulag1=nmu
! Update Open, Close, High and Low prices
      IF (n.EQ.1) THEN
       po(n)=100.0D0
       npo(n)=100.0D0
      ELSE
       po(n)=pc(n-1)
       npo(n)=npc(n-1)
      ENDIF
      pc(n)=po(n)*DEXP(y(steps))
      ph(n)=po(n)*DEXP(yhi)
      pl(n)=po(n)*DEXP(ylo)
      npc(n)=npo(n)*DEXP(ny(steps))
      nph(n)=npo(n)*DEXP(nyhi)
      npl(n)=npo(n)*DEXP(nylo)
     ENDDO
! Day-Loop; End
! Output
     OPEN(9,FILE=BIGfile//'.txt',STATUS='Replace')
     OPEN(10,FILE=BMfile//'.txt',STATUS='Replace')
     WRITE(9,*) 'Y RV HIGH LOW W'
     WRITE(10,*) 'Y RV HIGH LOW W'
     DO t=1,tt
! Write Log-returns, Realised Volatilities and Random Impact Factors to Files
      WRITE(9,'(5F17.10)') DLOG(pc(t))-DLOG(po(t)),rv(t),
    &  DLOG(ph(t))-DLOG(po(t)),DLOG(po(t))-DLOG(pl(t)),w(t)
      WRITE(10,'(5F17.10)') DLOG(npc(t))-DLOG(npo(t)),nrv(t),
    &  DLOG(nph(t))-DLOG(npo(t)),DLOG(npo(t))-DLOG(npl(t)),1.0D0
     ENDDO
     CLOSE(9)
     CLOSE(10)
     WRITE(6,*) tt,'Y RV HIGH LOW W (ID-BIG/TD) Written to '
    &    BIGfile//'.txt'
     WRITE(6,*) tt,'Y RV HIGH LOW W (ID-BM) Written to '
    &    BMfile//'.txt'
! De-allocate arrays
     DEALLOCATE(w,b,y,ny)
     RETURN
     END SUBROUTINE GEN_BM_and_BIG_or_TD
!=======================================================================
```

A.3.2 Log-likelihood Function of the ID-BIG Model

The following subroutine calculates the log-likelihood function of the ID-BIG model of Section 3.2.3. The module is an excerpt from the actual module, and contains only the global variables relevant for this subroutine. The residuals and estimates calculated at the optimal parameter values are later used to test the model fit by comparing the residuals and estimates to their respective model implied distributions using Q-Q and PIT plots, as reported on in Section 3.3. The variables $y(t)$ and $r(t)$ contain the daily returns and realised volatilities respectively.

```
! (Excerpt of) Module for global variables
!=======================================================================
      MODULE modA
!=======================================================================
      INTEGER(4) t,tt,maxn,skip,steps
      CHARACTER*(*) root
      LOGICAL mesg,resd
      PARAMETER (tt=1050,maxn=5050,skip=50,steps=78,root='C:\PhDFortran')
      REAL(8) y(maxn),r(maxn),mu(maxn),h(maxn), erz(maxn),erx(maxn),
     &        erv(maxn),ert(maxn),erw(maxn),erc(maxn),hth(maxn),htm(maxn)
      END MODULE
!=======================================================================

! ID-BIG Model Log-likelihood Function.
! X[1..7]=nu,ar1,arch0,arch1,garch1,beta,psi
!=======================================================================
      SUBROUTINE ML_ID-BIG(NX,X,LogLik)
!=======================================================================
! Declarations
      USE modA
      IMPLICIT NONE
      INTEGER(4) NX,stp
      REAL(8) X(NX),LogLik,som,Dens_SNCIG,st,isn,xn,vn,del,gam,p2,p3,p4,
     &        pp,bbb,mybesk
      EXTERNAL Dens_SNCIG,mybesk ! mybesk=Bessel function of 3rd order

! Initiate variables
      som=0.0D0; st=steps; stp=steps-1; pp=x(7)*x(7); gam=DSQRT(pp+x(6)*x(6))
! Day Loop; Start
      DO t=1,tt
! AR(1)-GARCH(1,1)
       IF (t.EQ.1) THEN
        mu(t)=X(1)*(1.0D0+X(2))
        h(t)=X(3)*(1.0D0-X(4))/(1.0D0-X(4)-X(5))
       ELSE
        h(t)=X(3)+X(4)*(y(predi+t-1)-mu(t-1))**2.0D0+X(5)*h(t-1)
        mu(t)=X(1)+X(2)*y(predi+t-1)
       ENDIF
       IF (h(t).LT.1.0D-10) h(t)=1.0D-10 ! Safety
! Log-likelihood part for day t
       IF (t.GT.skip) THEN ! Skip days to stabilise mu and h
        isn=MAX(DFLOAT(steps)*r(t)-y(t)*y(t),1.0D-200) ! Safety
        vn=isn/h(t); xn=(y(t)-mu(t))/DSQRT(h(t))
        som=som+1.5D0*DLOG(h(t))-
     &       DLOG(MAX(Dens_SNCIG(x(6),x(7),stp,xn,vn),1.0D-150)) ! Safety
! Calculate residuals (resd=.TRUE. only after optimal fit obtained)
        IF (resd) THEN
         del=DSQRT(pp+xn*xn+vn); p2=del*gam; p3=DSQRT(gam/del); p4=del/gam
         bbb=mybesk(p2,0.5D0*(stp+2.0D0))
         erw(t)=p4*mybesk(p2,0.5D0*stp)/bbb
         hth(t)=h(t); htm(t)=mu(t); erx(t)=xn; erv(t)=vn
         ert(t)=DSQRT((st-1.0D0)/isn)*(y(t)-mu(t))
         erz(t)=xn*p3*mybesk(p2,0.5D0*(stp+3.0D0))/bbb-
     &          x(6)*mybesk(p2,0.5D0*(stp+1.0D0))/(p3*bbb)
         erc(t)=vn*mybesk(p2,0.5D0*(stp+4.0D0))/(p4*bbb)
        ENDIF ! resd
```

```
       ENDIF ! t>skip
      ENDDO !t
! Day-Loop; End
! Log-likelihood
      LogLik=som-(tt-skip)*DLOG(st)
      IF (mesg) THEN
       write(6,'(7F9.5,F12.5)') X(1),x(2),x(3)*1.0D5,x(4),x(5),X(6),X(7),
     &             LogLik/(tt-skip)
      ENDIF
      RETURN
      END SUBROUTINE ML_ID-BIG
!=======================================================================
```

A.3.3 Fitting the TD-ID-BIG Model

We now take an in-depth look at the program that fits the TD-ID-BIG model of Section 4.2. We will concentrate our discussion on the maximisation of the log-likelihood function, while the more basic parts of the program are included for reference.

Firstly, the returns and realised volatilities are read from the input file. We then fit the ID-BIG model to the data to obtain a starting point for the TD-ID-BIG model, which we fit subsequently. When calculating the log-likelihood for the TD-ID-BIG model, three separate log-likelihood functions are actually calculated, namely:

- The ordinary ID-BIG log-likelihood where we assume $\rho = 0$.
- The TD-ID-BIG log-likelihood based on EIS where we assume $\rho = 0$. We implement the control variate variance reduction techniques discussed in Section 4.2, where we assume the $t_{0,m}$'s are normally distributed.
- The TD-ID-BIG log-likelihood based on EIS where $\rho \neq 0$, again implementing the control variate variance reduction techniques described in Section 4.2, assuming the t_m's are normally distributed.

These values are then combined to provide a more efficient (with smaller variance) single TD-ID-BIG log-likelihood function estimate (also as described in Section 4.2). It should be noted that common random numbers are used ("zz" and "zz2" in the program below) when calculating the TD-ID-BIG log-likelihood estimate, so that the estimate depends on the values of the parameters only, and not on a different set of random numbers used with each function call. Once the optimal parameter values (and their standard errors) have been obtained we calculate the residuals and estimates for residual analysis, as explained in Section 3.2.3 and the end of Section 4.2. The distributions of the residuals and estimates are then compared to their model implied distributions and they are then written to the output file together with the values needed for the PIT plots (the values for the Q-Q plots can be derived from the values for the PIT plots).

```
! Module for global variables
!=======================================================================
      MODULE modA
!=======================================================================
      INTEGER(4) t,tt,mm,mm2,maxn,skip,iseed,pitn,steps
      REAL(8) large,pi,nu,ar1,arch0,arch1,garch1,bta,ps,rho
      CHARACTER*(*) root
      CHARACTER*(*) input,output,outpar
      LOGICAL mesg,resd
      REAL(8),DIMENSION(:,:),ALLOCATABLE::UU,UU0,XX,ZZ,BB,ZZ2,CC
      REAL(8),DIMENSION(:),ALLOCATABLE::YY,ewxv,ecxv,ezxv,etxv
      PARAMETER (pi=3.141592653589793D0,maxn=5050,large=50000.0D0,
     &           pitn=500000,root='c:\PhDFortran')
      REAL(8) y(maxn),rv(maxn),mu(maxn),h(maxn),ymin,ymax,ymean,tmp(maxn),
     &         erx(maxn),hth(maxn),htm(maxn),htx(maxn),htw(maxn), htz(maxn),
     &         erz(maxn),erw(maxn),erv(maxn), ert(maxn),erc(maxn),htv(maxn),
     &         htt(maxn),htc(maxn),erw2(maxn),htw2(maxn)
      INTEGER(4) pitw(maxn),pitc(maxn),pitt(maxn),pitz(maxn),pitw2(maxn)
      END MODULE
!=======================================================================

! Program to fit the TD-ID-BIG model to Returns and Realised Volatilities
! Output: a) Optimal parameter values & Standard errors
!         b) Values for PIT plots
!=======================================================================
      PROGRAM TD-ID-BIG
!=======================================================================
! Declarations
      USE MSIMSL
      USE modA
      IMPLICIT NONE

! Initiate variables
      mesg=.TRUE.; resd=.FALSE.; steps=78; skip=50 ! To stabilise
      iseed=112233; CALL RNSET(iseed); mm=150; mm2=200
      input=root//'\TD-ID-BIGGen5.txt'
      output=root//'\TD-ID-BIGGen5Err.txt'
      outpar=root//'\TD-ID-BIGGen5Par.txt'
      WRITE(6,*) 'Input data: '//input
      WRITE(6,*) 'Output data: '//output//' and '//outpar
      WRITE(6,*) 'Using ',steps, 'steps.'
      OPEN(3,FILE=Output,STATUS='Unknown')
      OPEN(4,FILE=OutPar,STATUS='Unknown')
! Read Log-returns and Realised volatilities into Y and RV
      CALL READ_DATA(tt,input)
      write(6,*) 'Data days used:',tt
! Allocate arrays
      ALLOCATE(uu(tt,max(mm,mm2)),zz(tt,mm),xx(mm,2),bb(tt,3),yy(mm),
     &         zz2(tt,mm2),uu0(tt,max(mm,mm2)),cc(tt,3))
! Setup normal Common Random Numbers to use in ML function
      DO t=1,tt
       CALL DRNNOR(mm,zz(t,:))
       CALL DRNNOR(mm2,zz2(t,:))
      ENDDO !t
      CALL RNGET(iseed)
! Fit TD-ID-BIG model to Y and RV data; Calculate standard errors
      CALL FIT_TD-ID-BIG
```

```
      mesg=.FALSE.
! Calculate standard errors for optimal parameter values
      CALL DepBIGCOVARIANCE(tt-skip,nu,ar1,arch0,arch1,garch1,bta,ps,rho)
! Calculations for PIT plots (checking model fit)
      CALL PIT_Output

! De-allocate arrays, Close files and End program
      DEALLOCATE(UU,ZZ,XX,BB,YY,ZZ2,UU0,CC))
      CLOSE(3)
      CLOSE(4)
      STOP
      CONTAINS
!=======================================================================

! Read data for tot days. Set tot=0 to read all available
! Columns: log-return, realised volatility
!=======================================================================
      SUBROUTINE READ_DATA (tot,input)
!=======================================================================
      INTEGER(4) I,tot !tot=0 for all
      Character*80 Input

      WRITE(6,*) 'Reading from: ',input
      OPEN(2,FILE=Input,STATUS='Old')
      t=0.0D0
      DO WHILE (.NOT.EOF(2))
       t=t+1
       READ(2,*) y(t),rv(t)
       IF (t.EQ.1) THEN
        ymin=y(t); ymax=y(t)
       ELSE
        IF (y(t).LT.ymin) ymin=y(t); IF (y(t).GT.ymax) ymax=y(t)
       ENDIF
       ymean=ymean+y(t)
       IF (t.EQ.tot) GOTO 90;
      ENDDO
90    ymean=ymean/t
      tot=t
      IF (t.EQ.0) THEN
       WRITE (6,*) 'File is empty!'
      ELSE
       WRITE(6,*) 'Number of records read from file:',t
      ENDIF
      CLOSE (2)
      RETURN
      END SUBROUTINE READ_DATA
!=======================================================================

! Fit TD-ID-BIG model to data
!=======================================================================
      SUBROUTINE FIT_TD-ID-BIG
!=======================================================================
      USE modA
      IMPLICIT NONE
      INTEGER(4) LDA,NCON,NEQ,NVAR,N,IACT(30),NACT,NOUT,MAXFCN,MAXLAG,
     &           IPRINT,ISEOPT,IMEAN,INFO
      PARAMETER (IPRINT=0,ISEOPT=0,IMEAN=1,MAXLAG=4, NCON=1,LDA=NCON,NEQ=0,
     &           NVAR=8)
```

```
      REAL(8) A2(NCON,NVAR),ACC,ALAMDA(NVAR),B2(NCON),OBJ,SOL(NVAR),
     &        XGUESS(NVAR),XLB(NVAR),XUB(NVAR),FVALUE,FVALUEMIN, ACV(0:MAXLAG),
     &        AC(0:MAXLAG),SEAC(MAXLAG),FTOL,xmean,sum1,sum2
      EXTERNAL ML_TD-ID-BIG
      DATA FTOL/1.0D-7/,ACC/0.0D0/

! Use ID-BIG fit as starting point
! FIT_ID-BIG subroutine not included here. It is very similar to this procedure,
! but using the ID-BIG log-likelihood subroutine of Section A.3.2.
      CALL FIT_ID-BIG
      XGUESS(1)=nu; XGUESS(2)=ar1; XGUESS(3)=arch0; XGUESS(4)=arch1;
      XGUESS(5)=garch1; XGUESS(6)=bta; XGUESS(7)=ps; XGUESS(8)=0.0D0
! Parameter constraints
      A2(1,1)=0.0D0; A2(1,2)=0.0D0; A2(1,3)=0.0D0; A2(1,4)=1.0D0
      A2(1,5)=1.0D0; A2(1,6)=0.0D0; A2(1,7)=0.0D0; A2(1,8)=0.0D0
      B2(1)=1.0D0-FTOL*1.0D1
      XLB(1)=Ymin; XUB(1)=Ymax
      XLB(2)=FTOL-1.0D0; XUB(2)=1.0D0-FTOL
      XLB(3)=FTOL*1.0D1; XUB(3)=large
      XLB(4)=FTOL; XUB(4)=1.0D0-FTOL
      XLB(5)=0.0D0; XUB(5)=1.0D0-FTOL
      XLB(6)=-large; XUB(6)=large
      XLB(7)=0.5D0;XUB(7)=large
      XLB(8)=-0.99D0;XUB(8)=0.99D0
! Minimise minus-log-likelihood function
      MAXFCN=2000
      CALL DLCONF(ML_TD-ID-BIG,NVAR,NCON,NEQ,A2,LDA,B2,XLB,XUB,XGUESS,
     &            ACC,MAXFCN,SOL,FVALUE,NACT,IACT,ALAMDA)
! Output
      WRITE(6,*) 'Optimal Solution:'
      WRITE(6,*) SOL(1),SOL(2),SOL(3)*1.0D5,SOL(4),SOL(5),SOL(6),
     &           SOL(7),SOL(8)
      WRITE(4,*) 'Optimal Solution: nu ar1 arch0 arch1 garch1 bta psi rho'
      WRITE(4,*) SOL(1),SOL(2),SOL(3)*1.0D5,SOL(4),SOL(5),SOL(6),
     &           SOL(7),SOL(8)
      nu=SOL(1);ar1=SOL(2);arch0=SOL(3);arch1=SOL(4);garch1=SOL(5)
      bta=SOL(6);ps=SOL(7);rho=SOL(8)

      RETURN
      END SUBROUTINE FIT_TD-ID-BIG
!=======================================================================

! Calculate standard errors of parameters at given values
!=======================================================================
      SUBROUTINE COV_TD-ID-BIG(m,nup,arp1,archp0,archp1,garchp1,btap,psip,rhop)
!=======================================================================
      USE modA
      REAL(8) nup,arp1,archp0,archp1,garchp1,btap,psip,rhop
      INTEGER(4) i,j,k,m,LDFJAC,N,LDHES
      PARAMETER (N=8, LDHES=N)
      REAL(8) C(N),XSCALE(N),FJAC(M,N),FVALUE,HES(LDHES,N),EPSFCN,FC(M)
      EXTERNAL ML_TD-ID-BIG
! Initiate variables
      XC(1)=nup; XC(2)=arp1; XC(3)=archp0; XC(4)=archp1
      XC(5)=garchp1; XC(6)=btap; XC(7)=psip; XC(8)=rhop
      DO i=1,n
       XSCALE(i)=1.0D0
      enddo
```

```
      XSCALE(3)=1.0D5 ! Necessary to scale for arch0x10E5
      LDFJAC=m
! Evaluate the function at current point; Also, calculate residuals
      resd=.TRUE. ! Calculate residuals
      CALL ML_TD-ID-BIG (N,XC,FVALUE)
      resd=.FALSE.
! Get Hessian forward difference approximation
      EPSFCN=0.0D0
      WRITE(6,*) 'Calculating Hessian...'
      CALL DFDHES (ML_TD-ID-BIG,N,XC,XSCALE,FVALUE,EPSFCN,HES,LDHES)
      DO i=1,N-1
       DO j=I+1,N
        HES(i,j)=HES(j,i)
       ENDDO
      ENDDO
! Iverse Hessian
      CALL DLINRG(N,HES,N,HES,N)
! Output standard errors
      HES=(DFLOAT(m)/DFLOAT(m-N))*HES
      WRITE (4,*) 'Standard errors:'
      DO i=1,N
       IF (HES(i,i).LT.0.0D0) THEN
        WRITE(6,*) 'Rounding/Singularity error ... Standard error set to 0.'
        HES(i,i)=0.0D0
       ENDIF
       IF (i.EQ.3) THEN ! arch0x10E5
        WRITE(4,*) DSQRT(HES(i,i))*1.0D5
       ELSE
        WRITE(4,*) DSQRT(HES(i,i))
       ENDIF
      ENDDO
      END SUBROUTINE COV_TD-ID-BIG
!=======================================================================

      END PROGRAM TD-ID-BIG
!=======================================================================

! Bessel function of the third order, whole and half indices
!=======================================================================
      REAL(8) FUNCTION mybesk(xx,llp)
      USE modA,only: pi
      USE MSIMSL
      REAL(8) xx,b0,b1,bh,bb,mm,lp,llp
      INTEGER(4) m,lp2

      lp=DABS(llp)
! Extreme value
      IF (xx.GE.700.0D0) THEN
       mybesk=1.0D-304
       RETURN
      ENDIF
! Implement formula for whole and half index Bessel function
      bh=DSQRT(0.5D0*pi)/(DEXP(xx)*DSQRT(xx))
      lp2=lp
      IF (lp2.EQ.lp) THEN ! Whole index
       mm=0.0D0; b0=DBSK0(xx)/bh; b1=DBSK1(xx)/bh
      ELSE ! Half index
       mm=0.5D0; b0=1.0D0; b1=b0
```

```
      ENDIF
      IF (lp.EQ.0.0D0) THEN
       bb=b0
      ELSEIF (lp.LE.1.0D0) THEN
       bb=b1
      ENDIF
      DO m=2,lp+mm
       bb=2.0D0*(m-mm-1.0D0)*b1/xx+b0; b0=b1; b1=bb
      ENDDO
      mybesk=bb*bh
      END FUNCTION mybesk
!=======================================================================

! Density and Distribution Functions
!=======================================================================
! Standard normal density function
!=======================================================================
      REAL(8) FUNCTION Dens_N01(xx)
      USE modA, only: pi
      REAL(8) xx
      Dens_N01=DEXP(-0.5D0*xx*xx)/DSQRT(2.0D0*pi)
      END FUNCTION Dens_N01
! UIG density function
!=======================================================================
      REAL(8) FUNCTION Dens_UIG(psip,x)
      USE modA, only: pi
      REAL(8) psip,x,pp
      pp=psip*psip
      Dens_UIG=psip*x**(-1.5D0)*DEXP(-0.5D0*pp*(-2.0D0+x+1.0D0/x))/
     &          DSQRT(2.0D0*pi)
      END FUNCTION Dens_UIG
! UIG distribution function
!=======================================================================
      REAL(8) FUNCTION Dist_UIG(psip,x)
      USE MSIMSL
      REAL(8) psip,x,sqx,tmp1,tmp2
      sqx=DSQRT(x)
      tmp1=DNORDF(psip*(sqx-1.0D0/sqx))
      tmp2=DNORDF(psip*(-sqx-1.0D0/sqx))
      Dist_UIG=tmp1+DEXP(2.0D0*psip*psip)*tmp2
      END FUNCTION Dist_UIG
! Inverse of UIG distribution function
!=======================================================================
      REAL(8) FUNCTION DistInv_UIG(psip,z) ! z is standard normal
      USE MSIMSL
      REAL(8) wk,newwk,temp,FTOL,FTOL2,psip,z,u,Dist_UIG,Dens_UIG
      INTEGER(4) cnt
      EXTERNAL Dist_UIG,Dens_UIG
      data FTOL/1.0D-8/,FTOL2/1.0D-2/
      cnt=0
      u=DNORDF(z)
! Starting value more conservative in extreme cases, else crashes occur
      IF (z.EQ.3.0D0) THEN
       wk=6.0D0
      ELSEIF (z.LT.-3.0D0) THEN
        wk=0.05D0
      ELSE
       temp=DLOG(1.0D0+1.0D0/(psip*psip))
```

```
       wk=DEXP(DSQRT(temp)*z-0.5D0*temp)
      ENDIF
      IF (wk.LT.FTOL2) wk=FTOL2
! Obtain inverse through Newton-Rhapson method
      temp=1.0D10
      DO WHILE ((DABS(temp).GT.FTOL).AND.(cnt.LT.100))
       cnt=cnt+1
       temp=(Dist_UIG(psip,wk)-u)/Dens_UIG(psip,wk)
       newwk=wk-temp
! If a drastic jump occurs, limit it to a reasonable size
       DO WHILE ((newwk.LT.FTOL2).OR.(newwk.GT.100.0D0))
        newwk=wk+temp; temp=temp/2.0D0; newwk=wk-temp
       ENDDO !while newwk
       wk=newwk
      ENDDO !while DABS
      DistInv_UIG=wk
      END FUNCTION DistInv_UIG
! SCIG density function
!=====================================================================
      REAL(8) FUNCTION Dens_SCIG(psip,lp,v)
      USE modA
      REAL(8) psip,v,pv,mybesk,DGAMMA
      INTEGER(4) lp
      EXTERNAL mybesk
      pv=DSQRT(psip*psip+v)
      Dens_SCIG=((lp+3.0D0)/2.0D0)*DLOG(psip)+psip*psip+
     &  (lp/2.0D0-1.0D0)*DLOG(v)+DLOG(mybesk(psip*pv,(lp+1.0D0)/2.0D0))-
     &  DLOG((DSQRT(pi)*DGAMMA(lp/2.0D0)*2.0D0**((lp-1.0D0)/2.0D0)*
     &       pv**((lp+1.0D0)/2.0D0)))
      Dens_SCIG=DEXP(Dens_SCIG)
      END FUNCTION Dens_SCIG
! TIG distribution function
!=====================================================================
      REAL(8) FUNCTION Dens_TIG(btap,psp,lp,x)
      USE modA
      REAL(8) btap,psp,x,FTOL,lastterm,btmag,rfac,psm,psdmag,DTDF,tempTIG,term,
     &        tweemag,lmag,tlmag,bh,bh0,bh1,b0,b1,bes,DBSK0,DBSK1,DGAMMA
      INTEGER(4) rr,lp,rmin
      LOGICAL even
      DATA FTOL /1.0D-15/
! t-distribution if btap=0
      IF (btap.EQ.0.0D0) THEN
       Dens_TIG=DTDF(x,lp)
       RETURN
      ENDIF
! Different upper bounds for different parameter value combinations
      IF ((DABS(btap).LT.0.5D0).AND.(psp.LT.3.0D0)) THEN
       rmin=50
      ELSEIF ((DABS(btap).LT.3.0D0).AND.(psp.LT.5.0D0)) THEN
       rmin=100
      ELSE
       rmin=200
      ENDIF
! Initiate variables
! Use recursive Bessel formulae here
      lastterm=FTOL+1.0D0; rr=0; rfac=1.0D0; term=0.0D0
      psm=psp*DSQRT(psp**2.0D0+btap**2.0D0)
      psdmag=0.5D0*DLOG(psp/DSQRT(psp**2.0D0+btap**2.0D0))
```

```
      tweemag=0.5D0*DLOG(2.0D0); btmag=1.0D0
      lmag=0.5D0*DLOG(DFLOAT(lp)); tlmag=0.5D0*DLOG(1.0D0+x**2.0D0/lp)
      bh=DSQRT(pi/2.0D0)*DEXP(-psm)/DSQRT(psm)
      b0=DBSK0(psm)/bh; b1=DBSK1(psm)/bh; bh0=1.0D0; bh1=1.0D0
      tempTIG=psp**2.0D0+DLOG(psp/pi); even=.TRUE.
! Repeat until required accuracy obtained
      DO WHILE ((lastterm.GT.FTOL).OR.(rr.LT.rmin))
! Cover all cases of Bessel functions
       IF (even) THEN
        IF (rr.LT.4) THEN
         bes=bh
        ELSE
! get Bessel n+1/2 from n-1/2 and n-3/2
         bes=(rr-3.0D0)*bh0/psm+bh1
         bh0=bh1
         bh1=bes
         bes=bes*bh
        ENDIF
       ELSE
! Get Bessel n from n-1 and n-2
        IF (rr.EQ.1) bes=b0
        IF (rr.EQ.3) bes=b1
        IF (rr.GE.5) THEN
         bes=(rr-3.0D0)*b1/psm+b0
         b0=b1
         b1=bes
        ENDIF
        bes=bes*bh
       ENDIF
! Calculate current term to be added
      IF (bes.LT.1.0D-100) THEN
       lastterm=FTOL
      ELSE
       lastterm=(rr+1.0D0)*tweemag+DLOG(DGAMMA((lp+rr+1.0D0)/
     &      2.0D0))+(rr-1.0D0)*psdmag-(rr+1.0D0)*lmag-DLOG(rfac)
     &      -DLOG(DGAMMA(lp/2.0D0))-(lp+rr+1)*tlmag+DLOG(bes)
       lastterm=DEXP(lastterm)*btmag
      ENDIF
       term=term+lastterm
       even=.NOT.even
       rr=rr+1
       rfac=rfac*rr
       btmag=btmag*x*btap
      ENDDO ! While
      IF (term.LT.1.0D-100) term=1.0D-100; ! Safety in extreme cases
      tempTIG=tempTIG+DLOG(term); Dens_TIG=DEXP(tempTIG)
      END FUNCTION Dens_TIG
! SNIG density function
!=======================================================================
      REAL(8) FUNCTION Dens_SNIG(btap,psip,x)
      USE modA
      REAL(8) btap,psip,x,bp,xp
      bp=DSQRT(btap*btap+psip*psip); xp=DSQRT(x*x+psip*psip)
      Dens_SNIG=psip*bp*DEXP(psip*psip+btap*x)*DBSK1(bp*xp)/(pi*xp)
      END FUNCTION Dens_SNIG
! SNCIG distribution function
!=======================================================================
      REAL(8) FUNCTION Dens_SNCIG(btap,psip,lp,x,v)
```

```
      USE modA, only: pi
      INTEGER(4) m,lp
      REAL(8) btap,psip,x,v,bp,xvp,lop2,b0,b1,bh,bb,xx,mm,mybesk,lp2,dgamma
      EXTERNAL mybesk
      bp=DSQRT(btap*btap+psip*psip); xvp=DSQRT(x*x+v+psip*psip); xx=bp*xvp
      lop2=0.5D0*lp
      IF (MOD(lp,2).EQ.0) THEN
       mm=0.0D0
      ELSE
       mm=0.5D0
      ENDIF
      lp2=lop2+1.0D0; bb=mybesk(xx,lp2)
      Dens_SNCIG=psip*(bp/xvp)**(lop2+1)*v**(lop2-1)*
     &    DEXP(max(min(psip*psip+btap*x,1.0D150,500.0D0),-500.0D0))*bb
     &   /(pi*DGAMMA(lop2)*(2.0D0**lop2))
      END FUNCTION Dens_SNCIG
!=======================================================================

! TD-ID-BIG Model Log-likelihood Function.
! X[1..8]=nu,ar1,arch0,arch1,garch1,beta,psi,rho
!=======================================================================
      SUBROUTINE ML_TD-ID-BIG(NX,X,AF)
!=======================================================================
      USE modA
      REAL(8) sig2,mut,lnchitp1,gu,temp,som,somi(max(mm,mm2)),sig20,ssig2,
     &        ssig20,mut0,lnchitp10,som0,somi0(max(mm,mm2)),som00,
     &        X(NX),SSE,SST,Dens_N01,DistInv_UIG, AF, Dens_SNCIG,
     &        coeftol,diffcoef,rho1,rho2,rho3,rho4,logrho1,stmin1,
     &        sttmp,log2,vn,dlngam,st,isn,xn,del,gam,p2,p3,p4,pp,bbb,mybesk
      INTEGER(4) NX,INTCEP, LDX, NCOEF, NIND, NOBS, totcnt,m,maxcnt,r,stp
      PARAMETER (INTCEP=1, NIND=2, NCOEF=INTCEP+NIND)
      REAL(8) C(NCOEF)
      DATA coeftol/1.0D-5/
      EXTERNAL Dens_N01,Dens_SNCIG,DistInv_UIG,mybesk
! Initiate variables
      pp=x(7)*x(7); gam=DSQRT(pp+x(6)*x(6))
      LDX=mm; rho1=DSQRT(1-X(8)*x(8)); rho2=1.0D0/(1.0D0-x(8)*x(8))
      rho3=x(8)*rho2; rho4=x(8)*rho3; logrho1=DLOG(rho1)
      Som00=0.0D0; st=steps; stp=steps-1
! Calculate exact log-likelihood at rho=0 (as in ID-BIG)
      DO t=1,tt
       IF (t.EQ.1) THEN
        mu(t)=X(1)*(1.0D0+X(2))
        h(t)=X(3)*(1.0D0-X(4))/(1.0D0-X(4)-X(5))
       ELSE
        h(t)=X(3)+X(4)*(y(t-1)-mu(t-1))**2.0D0+X(5)*h(t-1)
        mu(t)=X(1)+X(2)*y(t-1)
       ENDIF
      IF (t.GT.skip) THEN
        isn=MAX(DFLOAT(steps)*rv(t)-y(t)*y(t),1.0D-200) ! Safety when few steps
        vn=isn/h(t); xn=(y(t)-mu(t))/DSQRT(h(t))
        Som00=Som00+1.5D0*DLOG(h(t))-
     &        DLOG(MAX(Dens_SNCIG(x(6),x(7),stp,xn,vn),1.0D-150))
       ENDIF
      ENDDO !t
      som00=som00-(tt-skip)*DLOG(st)
! Start; Calculate log-likelihood for rho=0 and rho≠0 using EIS
! For rho=0 ...
```

```
      totcnt=1; diffcoef=1.0D5; maxcnt=40.0D0
      stmin1=0.5D0*(steps-1.0D0); log2=stmin1*DLOG(2.0D0)+DLNGAM(stmin1)
      DO WHILE ((totcnt.LE.maxcnt).AND.(diffcoef.GT.coeftol))
       IF (totcnt.EQ.1) THEN
! rho=0 u sampled from starting p
        DO m=1,mm
         uu(1,m)=zz(1,m)
         DO t=2,tt
          uu(t,m)=zz(t,m)
         ENDDO !t
        ENDDO !m
       ELSE
! rho=0 u sampled from calculated q
        DO m=1,mm
         sig2=1.0D0/(1.0D0-2.0D0*bb(1,3)) ! t=1 case
         mut=sig2*bb(1,2)
         uu(1,m)=DSQRT(sig2)*zz(1,m)+mut
         DO t=2,tt
          sig2=1.0D0/(1.0D0-2.0D0*bb(t,3))
          mut=sig2*bb(t,2)
          uu(t,m)=DSQRT(sig2)*zz(t,m)+mut
         ENDDO !t
        ENDDO !m
       ENDIF
       IF (totcnt.GT.1) diffcoef=0.0D0
       DO t=tt,skip+1,-1
        DO m=1,mm
         xx(m,1)=uu(t,m); xx(m,2)=uu(t,m)*uu(t,m)
! rho=0 EIS calculations
         IF (t.EQ.tt) THEN
          lnchitp1=0.0D0
         ELSE
          sig2=1.0D0/(1.0D0-2.0D0*bb(t+1,3))
          mut=sig2*bb(t+1,2)
          lnchitp1=0.5D0*DLOG(sig2)+bb(t+1,1)+0.5D0*mut*mut/sig2
         ENDIF
         gu=DistInv_UIG(x(7),uu(t,m))
         vn=(steps*rv(t)-y(t)*y(t))/(h(t)*gu)
         IF (vn.LT.1.0D-5) vn=1.0D-5 ! Safety when few steps
         sttmp=(stmin1-1.0D0)*DLOG(vn)-0.5D0*vn-log2
         temp=(y(t)-mu(t)-x(6)*DSQRT(h(t))*gu)/DSQRT(h(t)*gu)
         temp=DLOG(Dens_N01(temp))-0.5D0*DLOG(h(t))-1.5D0*DLOG(gu)+sttmp
         yy(m)=temp+lnchitp1
        ENDDO !m
! rho=0 Linear regression to obtain the constants
        CALL DRLSE (mm, YY, NIND, XX, LDX, INTCEP, C, SST, SSE)
        IF (totcnt.GT.1) THEN
         diffcoef=MAX(diffcoef,DABS(bb(t,1)-C(1)),DABS(bb(t,2)-C(2)),
     &          DABS(bb(t,3)-C(3)))
        ELSE
         diffcoef=1.0D5
        ENDIF
        bb(t,:)=C
       ENDDO !t
       totcnt=totcnt+1
      ENDDO !while totcnt
! rho=0 Save constants of fit for later use
      cc=bb
```

```
! For rho≠0 ...
      totcnt=1; diffcoef=1.0D5
      DO WHILE ((totcnt.LE.maxcnt).AND.(diffcoef.GT.coeftol))
! rho≠0 u sampled from calculated q (last q of rho=0 case)
        DO m=1,mm
         sig2=1.0D0/(1.0D0-2.0D0*bb(1,3)) ! t=1 case
         mut=sig2*bb(1,2)
         uu(1,m)=DSQRT(sig2)*zz(1,m)+mut
         DO t=2,tt
          sig2=1.0D0/(rho2-2.0D0*bb(t,3))
          mut=sig2*(bb(t,2)+rho3*uu(t-1,m))
          uu(t,m)=DSQRT(sig2)*zz(t,m)+mut
         ENDDO !t
        ENDDO !m
       IF (totcnt.GT.1) diffcoef=0.0D0
       DO t=tt,skip+1,-1
        DO m=1,mm
         xx(m,1)=uu(t,m); xx(m,2)=uu(t,m)*uu(t,m)
! rho≠0 EIS calculations
         IF (t.EQ.tt) THEN
          lnchitp1=0.0D0
         ELSE
          sig2=1.0D0/(rho2-2.0D0*bb(t+1,3))
          mut=sig2*(bb(t+1,2)+rho3*uu(t,m))
          lnchitp1=0.5D0*DLOG(sig2)-logrho1+bb(t+1,1)+0.5D0*mut*mut/sig2-
     &      0.5D0*rho4*uu(t,m)*uu(t,m)
         ENDIF
         gu=DistInv_UIG(x(7),uu(t,m))
         vn=(steps*rv(t)-y(t)*y(t))/(h(t)*gu)
         IF (vn.LT.1.0D-5) vn=1.0D-5 ! Safety when few steps
         sttmp=(stmin1-1.0D0)*DLOG(vn)-0.5D0*vn-log2
         temp=(y(t)-mu(t)-x(6)*DSQRT(h(t))*gu)/DSQRT(h(t)*gu)
         temp=DLOG(Dens_N01(temp))-0.5D0*DLOG(h(t))-1.5D0*DLOG(gu)+sttmp
         yy(m)=temp+lnchitp1
        ENDDO !m
! rho≠0 Linear regression to obtain the constants
        CALL DRLSE (mm, YY, NIND, XX, LDX, INTCEP, C, SST, SSE)
        IF (totcnt.GT.1) THEN
         diffcoef=MAX(diffcoef,DABS(bb(t,1)-C(1)),DABS(bb(t,2)-C(2)),
     &         DABS(bb(t,3)-C(3)))
        ELSE
         diffcoef=1.0D5
        ENDIF
        bb(t,:)=C
       ENDDO !t
       totcnt=totcnt+1
      ENDDO !while totcnt
! Final choice of uu-distributions obtained
! Now calculate EIS log-likelihood for rho=0 and rho≠0
! This part is very similar to the part above
      sig2=1.0D0/(1.0D0-2.0D0*bb(1,3)); ssig2=DSQRT(sig2)! t=1 case
      sig20=1.0D0/(1.0D0-2.0D0*cc(1,3)); ssig20=DSQRT(sig20)
      mut=sig2*bb(1,2); mut0=sig20*cc(1,2)
      DO m=1,mm2;
       uu(1,m)=ssig2*zz2(1,m)+mut
       uu0(1,m)=ssig20*zz2(1,m)+mut
      ENDDO
```

```
      DO t=2,tt
       sig2=1.0D0/(rho2-2.0D0*bb(t,3)); ssig2=DSQRT(sig2)
       sig20=1.0D0/(1.0D0-2.0D0*cc(t,3)); ssig20=DSQRT(sig20)
       DO m=1,mm2
        mut=sig2*(bb(t,2)+rho3*uu(t-1,m)); mut0=sig20*cc(t,2)
        uu(t,m)=ssig2*zz2(t,m)+mut; uu0(t,m)=ssig20*zz2(t,m)+mut0
       ENDDO !m
      ENDDO !t
      somi=0.0D0; somi0=0.0D0
      DO m=1,mm2
       DO t=tt,skip,-1
         IF (t.EQ.tt) THEN
          lnchitp1=0.0D0; lnchitp10=0.0D0
         ELSE
          sig2=1.0D0/(rho2-2.0D0*bb(t+1,3))
          sig20=1.0D0/(1.0D0-2.0D0*cc(t+1,3))
          mut=sig2*(bb(t+1,2)+rho3*uu(t,m))
          mut0=sig20*cc(t+1,2)
          lnchitp1=0.5D0*DLOG(sig2)-logrho1+bb(t+1,1)+0.5D0*mut*mut/sig2-
     &       0.5D0*rho4*uu(t,m)*uu(t,m)
          lnchitp10=0.5D0*DLOG(sig20)+cc(t+1,1)+0.5D0*mut0*mut0/sig20
         ENDIF
         IF (t.GT.skip) THEN
          gu=DistInv_UIG(x(7),uu(t,m))
          vn=(steps*rv(t)-y(t)*y(t))/(h(t)*gu)
          IF (vn.LT.1.0D-5) vn=1.0D-5 ! Safety when few steps
          sttmp=(stmin1-1.0D0)*DLOG(vn)-0.5D0*vn-log2
          temp=(y(t)-mu(t)-x(6)*DSQRT(h(t))*gu)/DSQRT(h(t)*gu)
          temp=DLOG(Dens_N01(temp))-0.5D0*DLOG(h(t))-1.5D0*DLOG(gu)+sttmp
          somi(m)=somi(m)+temp+lnchitp1-bb(t,1)-bb(t,2)*uu(t,m)-
     &                  bb(t,3)*uu(t,m)*uu(t,m)
          gu=DistInv_UIG(x(7),uu0(t,m))
          vn=(steps*rv(t)-y(t)*y(t))/(h(t)*gu)
          if (vn.LT.1.0D-5) vn=1.0D-5 ! Safety when few steps
          sttmp=(stmin1-1.0D0)*DLOG(vn)-0.5D0*vn-log2
          temp=(y(t)-mu(t)-x(6)*DSQRT(h(t))*gu)/DSQRT(h(t)*gu)
          temp=DLOG(Dens_N01(temp))-0.5D0*DLOG(h(t))-1.5D0*DLOG(gu)+sttmp
          somi0(m)=somi0(m)+temp+lnchitp10-cc(t,1)-
     &               cc(t,2)*uu0(t,m)-cc(t,3)*uu0(t,m)*uu0(t,m)
         ENDIF
       ENDDO !t
       somi(m)=somi(m)+lnchitp1; somi0(m)=somi0(m)+lnchitp10
      ENDDO !m
      som=0.0D0; som0=som
      IF (resd) THEN
       WRITE(4,*) 'Ym Y0m'
       DO m=1,mm2
        WRITE(4,*) somi(m),somi0(m)
       ENDDO
      ENDIF !resd
! Calculate mean and variance of tm's (somi) and t0m's (som0i)
      mut=0; DO m=1,mm2; mut=mut+somi(m); ENDDO; mut=mut/mm2;
      sig2=0; do m=1,mm2;sig2=sig2+(somi(m)-mut)**2.0D0; ENDDO
      sig2=sig2/mm2; som=-mut-0.5D0*sig2
      mut=0; DO m=1,mm2; mut=mut+somi0(m); ENDDO; mut=mut/mm2;
      sig2=0; DO m=1,mm2;sig2=sig2+(somi0(m)-mut)**2.0D0; ENDDO
      sig2=sig2/mm2; som0=-mut-0.5D0*sig2
```

```
! Variance-reduced log-likelihood
! som and som0 are computed assuming the tm's and t0m's are normally distributed
! som00 is the exact ID-BIG log-likelihood function
! Now, add everything together for the logLnor estimate of the log-likelihood
      AF=som+som00-som0
! End; Log-likelihood obtained
! Calculate residuals and estimates (only performed on optimal solution)
      IF (resd) THEN
! Calculate residuals and estimates given X,V (as in ID-BIG)
       DO r=skip+1,tt
        isn=MAX(DFLOAT(steps)*rv(r)-y(r)*y(r),1.0D-200)
        vn=isn/h(r); xn=(y(r)-mu(r))/DSQRT(h(r))
        IF (vn.LT.1.0D-5) vn=1.0D-5 ! Safety when few steps
        del=DSQRT(pp+xn*xn+vn); p2=del*gam; p4=del/gam; p3=DSQRT(gam/del)
        bbb=mybesk(p2,0.5D0*(stp+2.0D0))
        hth(r)=h(r); htm(r)=mu(r); erx(r)=xn; erv(r)=vn
        ert(r)=dsqrt((st-1.0D0)/isn)*(y(r)-mu(r))
        erz(r)=xn*p3*mybesk(p2,0.5D0*(stp+3.0D0))/bbb-
     &        x(6)*mybesk(p2,0.5D0*(stp+1.0D0))/(p3*bbb)
        erc(r)=vn*mybesk(p2,0.5D0*(stp+4.0D0))/(p4*bbb)
        erw(r)=p4*mybesk(p2,0.5D0*stp)/bbb
! Estimates for Wnor
! This part is exactly like the part calculating the log-likelihood above,
! except where temp is given a value, commented below
       totcnt=1; diffcoef=1.0D5
       DO WHILE ((totcnt.LE.maxcnt).AND.(diffcoef.GT.coeftol))
        DO m=1,mm
         sig2=1.0D0/(1.0D0-2.0D0*bb(1,3))
         mut=sig2*bb(1,2)
         uu(1,m)=DSQRT(sig2)*zz(1,m)+mut
         DO t=2,tt
          sig2=1.0D0/(rho2-2.0D0*bb(t,3))
          mut=sig2*(bb(t,2)+rho3*uu(t-1,m))
          uu(t,m)=DSQRT(sig2)*zz(t,m)+mut !*zzw
         ENDDO !t
        ENDDO !m
        IF (totcnt.GT.1) diffcoef=0.0D0
        DO t=tt,skip+1,-1
         DO m=1,mm
          xx(m,1)=uu(t,m); xx(m,2)=uu(t,m)*uu(t,m)
          IF (t.EQ.tt) THEN
           lnchitp1=0.0D0
          ELSE
           sig2=1.0D0/(rho2-2.0D0*bb(t+1,3))
           mut=sig2*(bb(t+1,2)+rho3*uu(t,m))
           lnchitp1=0.5D0*DLOG(sig2)-logrho1+bb(t+1,1)+0.5D0*mut*mut/sig2-
     &       0.5D0*rho4*uu(t,m)*uu(t,m)
          ENDIF
          gu=DistInv_UIG(x(7),uu(t,m))
          vn=(steps*rv(t)-y(t)*y(t))/(h(t)*gu)
          IF (vn.LT.1.0D-5) vn=1.0D-5 ! Safety when few steps
          sttmp=(stmin1-1.0D0)*DLOG(vn)-0.5D0*vn-log2
          temp=(y(t)-mu(t)-x(6)*DSQRT(h(t))*gu)/DSQRT(h(t)*gu)
          IF (t.EQ.r) THEN
! Difference: Multiplied by g(u), thus power(multiplier) one more
           temp=DLOG(Dens_N01(temp))-0.5D0*DLOG(h(t))-0.5D0*DLOG(gu)+sttmp
          ELSE
           temp=DLOG(Dens_N01(temp))-0.5D0*DLOG(h(t))-1.5D0*DLOG(gu)+sttmp
```

```
          ENDIF
          yy(m)=temp+lnchitp1
         ENDDO !m
         CALL DRLSE (mm, YY, NIND, XX, LDX, INTCEP, C, SST, SSE)
         IF (totcnt.GT.1) THEN
          diffcoef=MAX(diffcoef,DABS(bb(t,1)-C(1)),DABS(bb(t,2)-C(2)),
     &         DABS(bb(t,3)-C(3)))
         ELSE
          diffcoef=1.0D5
         ENDIF
         bb(t,:)=C
        ENDDO !t
        totcnt=totcnt+1
       ENDDO !while totcnt
! Final choice of uu-distribution obtained, now Wnor estimate
       sig2=1.0D0/(1.0D0-2.0D0*bb(1,3)); ssig2=DSQRT(sig2)! t=1 case
       mut=sig2*bb(1,2)
       DO m=1,mm2
        uu(1,m)=ssig2*zz2(1,m)+mut
       ENDDO
       DO t=2,tt
        sig2=1.0D0/(rho2-2.0D0*bb(t,3)); ssig2=DSQRT(sig2)
        DO m=1,mm2
         mut=sig2*(bb(t,2)+rho3*uu(t-1,m))
         uu(t,m)=ssig2*zz2(t,m)+mut
        ENDDO !m
       ENDDO !t
       somi=0.0D0
       DO m=1,mm2
        DO t=tt,skip,-1
         IF (t.EQ.tt) THEN
          lnchitp1=0.0D0
         ELSE
          sig2=1.0D0/(rho2-2.0D0*bb(t+1,3))
          mut=sig2*(bb(t+1,2)+rho3*uu(t,m))
          lnchitp1=0.5D0*DLOG(sig2)-logrho1+bb(t+1,1)+0.5D0*mut*mut/sig2-
     &         0.5D0*rho4*uu(t,m)*uu(t,m)
         ENDIF
         IF (t.GT.skip) THEN
          gu=DistInv_UIG(x(7),uu(t,m))
          vn=(steps*rv(t)-y(t)*y(t))/(h(t)*gu)
          if (vn.LT.1.0D-5) vn=1.0D-5 ! Safety when few steps
          sttmp=(stmin1-1.0D0)*DLOG(vn)-0.5D0*vn-log2
          temp=(y(t)-mu(t)-x(6)*DSQRT(h(t))*gu)/DSQRT(h(t)*gu)
          IF (t.EQ.r) THEN
! Difference: Multiplied by g(u), thus power(multiplier) one more
           temp=DLOG(Dens_N01(temp))-0.5D0*DLOG(h(t))-0.5D0*DLOG(gu)+sttmp
          ELSE
           temp=DLOG(Dens_N01(temp))-0.5D0*DLOG(h(t))-1.5D0*DLOG(gu)+sttmp
          ENDIF
          somi(m)=somi(m)+temp+lnchitp1-bb(t,1)-bb(t,2)*uu(t,m)-
     &                  bb(t,3)*uu(t,m)*uu(t,m)
         ENDIF
        ENDDO !t
        somi(m)=somi(m)+lnchitp1
       ENDDO !m
        mut=0; DO m=1,mm2; mut=mut+somi(m); ENDDO; mut=mut/mm2;
        sig2=0; DO m=1,mm2;sig2=sig2+(somi(m)-mut)**2.0D0; ENDDO
```

```
         sig2=sig2/mm2
         erw2(r)=dexp(mut+0.5D0*sig2+som)
        ENDDO !r
      ENDIF !resd
! Output
      IF (mesg) THEN
       WRITE(6,'(8F8.5,F12.8)') x(1),x(2),x(3)*1.0D5,x(4),x(5),
     &          x(6),x(7),x(8),af/(tt-skip)
      ENDIF
      END SUBROUTINE ML_TD-ID-BIG
!=======================================================================

! Generate independent UIG-distributed variables
!=======================================================================
      SUBROUTINE Gen_UIG(tt,v,ps,iseed)
      REAL(8) ps,sqps, nu,vv,v(tt),z2,z1
      INTEGER(4) t,iseed,tt
      REAL(8),DIMENSION(:),ALLOCATABLE:: un,ru
! Initiate variables and Allocate arrays
      sqps=ps*ps
      ALLOCATE(un(tt),ru(tt))
! Generate uniform and N(0,1) variables
      CALL RNSET(iseed)
      CALL DRNUN(tt,ru)
      CALL DRNNOR(tt,un)
      CALL RNGET(iseed)
! Implement formulae
      DO t=1,tt
       vv=un(t)**2.0D0
       z1=1+(vv-dsqrt(4.0D0*sqps*vv+vv**2.0D0))/(2.0D0*sqps)
       z2=1.0D0/z1
       IF (ru(t).LT.1.0D0/(1.0D0+z1)) THEN
         v(t)=z1
        ELSE
         v(t)=z2
       ENDIF
      ENDDO
! De-allocate arrays
      DEALLOCATE(un,ru)
      RETURN
      END SUBROUTINE Gen_UIG
!=======================================================================

! Empirical distribution simulation needed for PIT calculations - residual
! analysis for X,V residuals and Z,C,W,T estimates
!=======================================================================
      SUBROUTINE PIT_SetupXV(btap,psp,lp,gm)
      USE modA
      INTEGER(4) i,m,gm,lp
      REAL(8) btap,psp,bta2,ps2,x,v,del,gam,dopg,bbb,mybesk,p2,p3
      EXTERNAL mybesk
! Initiate variables and Allocate arrays
      WRITE(6,*) 'Generating ',gm,' Z,C,W,T empirical distributions...'
      bta2=btap*btap; ps2=psp*psp; gam=DSQRT(ps2+bta2)
      ALLOCATE(ewxv(gm),ecxv(gm),ezxv(gm),etxv(gm))
! Generate N(0,1) variables Z
      CALL RNSET(iseed)
      CALL DRNNOR(gm,ezxv)
```

```
! Generate Chi-squared(lp) variables C
      DO m=1,gm
       CALL DRNNOR(lp,ewxv); ecxv(m)=0.0D0
       DO i=1,lp
        ecxv(m)=ecxv(m)+ewxv(i)**2.0D0
       ENDDO !i
      ENDDO !m
! Generate UIG(ps) variables W
      CALL GEN_UIG(gm,ewxv,psp,ISEED)
      CALL RNGET(iseed)
! Calculate corresponding X,V residuals and Z,C,W,T estimates
      DO m=1,gm
       x=btap*ewxv(m)+DSQRT(ewxv(m))*ezxv(m); v=ewxv(m)*ecxv(m)
       del=DSQRT(ps2+x*x+v); p2=del*gam;dopg=del/gam; p3=DSQRT(gam/del)
       bbb=mybesk(p2,0.5D0*(lp+2.0D0))
       ezxv(m)=x*p3*mybesk(p2,0.5D0*(lp+3.0D0))/bbb-
     &          btap*mybesk(p2,0.5D0*(lp+1.0D0))/(p3*bbb)
       ecxv(m)=v*mybesk(p2,0.5D0*(lp+4.0D0))/(dopg*bbb)
       ewxv(m)=dopg*mybesk(p2,0.5D0*lp)/bbb
       etxv(m)=x/DSQRT(v/lp)
      ENDDO !m
! Sort Z,C,W,T estimates
      CALL DSVRGN(pitn,ewxv,ewxv)
      CALL DSVRGN(pitn,ecxv,ecxv)
      CALL DSVRGN(pitn,ezxv,ezxv)
      CALL DSVRGN(pitn,etxv,etxv)
      END SUBROUTINE PIT_SetupXV

! Numerical integration: When calling the DQDAG routine later on, the first
! parameter passed to it has to be the function to be integrated, and it is
! required to have only one input  parameter (x). It is therefore necessary to
! define the following functions to call the respective density functions
! that has to be integrated, but ensure that these functions are evaluated
! at the given x and the optimal parameters obtained from the ML fit. Note
! that the parameters bta, ps and steps are global variables and bta and ps
! will contain the optimal ML parameter fit by the time these functions are
! required.
!=======================================================================
! SNIG density function: version needed for numerical integration DQDAG
!=======================================================================
      REAL(8) FUNCTION Dens_SNIG_PIT(x)
      USE modA
      REAL(8) x
      Dens_SNIG_PIT=Dens_SNIG(bta,ps,x)
      END FUNCTION Dens_SNIG_PIT
! SCIG density function: version needed for numerical integration DQDAG
!=======================================================================
      REAL(8) FUNCTION Dens_SCIG_PIT(x)
      USE modA
      REAL(8) x
      Dens_SCIG_PIT=Dens_SCIG(MIN(20.0D0,ps),steps-1,x) ! Cap ps at 20
      END FUNCTION Dens_SCIG_PIT
!=======================================================================

! PIT calculations to evaulate model fit (residual analysis)
!=======================================================================
      SUBROUTINE PIT_Output
      USE modA
```

```
      USE MSIMSL
      INTEGER(4) IRULE,mw,mc,mz,mt,mw2
      REAL(8) a,b,res,err,err2
      EXTERNAL Dens_SNIG_PIT,Dens_SCIG_PIT,Dens_TIG_PIT,Dens_UIG_PIT
! Save residuals and estimates for output
      htz=erz;htx=erx;htw=erw;htw2=erw2
      htv=erv;htc=erc;htt=ert
      WRITE(6,*) 'Performing PIT calculations (residual anaylsis)...'
! Sort residuals and estimates
      CALL DSVRGN(tt-skip,erx(skip+1:tt),erx(skip+1:tt))
      CALL DSVRGN(tt-skip,erv(skip+1:tt),erv(skip+1:tt))
      CALL DSVRGN(tt-skip,ert(skip+1:tt),ert(skip+1:tt))
      CALL DSVRGN(tt-skip,erw(skip+1:tt),erw(skip+1:tt))
      CALL DSVRGN(tt-skip,erw2(skip+1:tt),erw2(skip+1:tt))
      CALL DSVRGN(tt-skip,erz(skip+1:tt),erz(skip+1:tt))
      CALL DSVRGN(tt-skip,erc(skip+1:tt),erc(skip+1:tt))
! Simulate empirical distributions at optimal ML parameter values
      CALL PIT_SetupXV(bta,ps,steps-1,PITN)
      IRULE=6; err=1.0D-10
      WRITE (3,*) 'y rv Unif X V T W Z C WXV h^ mu^ z^ x^ v^ t^ w^ c^ wxv^'
! Day-Loop; Start
      DO t=skip+1,tt
! Numerical integration for X and V residuals
       a=-500.0D0; b=erx(t)
       CALL DQDAG(Dens_SNIG_PIT,a,b,err,err,irule,erx(t),err2)
       a=1.0D-20; b=erv(t)
       CALL DQDAG(Dens_SCIG_PIT,a,b,err,err,irule,erv(t),err2)
! Compare Z,C,W,T estimates with simulated distributions
! Note the repetition for Z,C,W,T
       IF (t.EQ.skip+1) THEN
        pitz(t)=0; pitc(t)=0; pitw(t)=0; pitw2(t)=0; pitt(t)=0
       ELSE
        pitz(t)=pitz(t-1); pitc(t)=pitc(t-1)
        pitw(t)=pitw(t-1); pitw2(t)=pitw2(t-1); pitt(t)=pitt(t-1)
       ENDIF
       mz=pitz(t)+1
       IF (mz.LE.pitn) THEN
        a=ezxv(mz)
       ELSE
        a=ezxv(pitn)
       ENDIF
       DO WHILE ((mz.LE.pitn).AND.(erz(t).GT.a))
        mz=mz+1
        IF (mz.LE.pitn) THEN
         a=ezxv(mz)
        ELSE
         a=ezxv(pitn)
        ENDIF
       ENDDO !while
       pitz(t)=mz-1
       mc=pitc(t)+1
       IF (mc.LE.pitn) THEN
        a=ecxv(mc)
       ELSE
        a=ecxv(pitn)
       ENDIF
       DO WHILE ((mc.LE.pitn).AND.(erc(t).GT.a))
        mc=mc+1
```

```
 IF (mc.LE.pitn) THEN
  a=ecxv(mc)
 ELSE
  a=ecxv(pitn)
 ENDIF
ENDDO !while
pitc(t)=mc-1
mw=pitw(t)+1
IF (mw.LE.pitn) THEN
 a=ewxv(mw)
ELSE
 a=ewxv(pitn)
ENDIF
DO WHILE ((mw.LE.pitn).AND.(erw(t).GT.a))
 mw=mw+1
 IF (mw.LE.pitn) THEN
  a=ewxv(mw)
 ELSE
  a=ewxv(pitn)
 ENDIF
ENDDO !while
pitw(t)=mw-1
mw2=pitw2(t)+1
IF (mw2.LE.pitn) THEN
 a=ewxv(mw2)
ELSE
 a=ewxv(pitn)
ENDIF
DO WHILE ((mw2.LE.pitn).AND.(erw2(t).GT.a))
 mw2=mw2+1
 IF (mw2.LE.pitn) THEN
  a=ewxv(mw2)
 ELSE
  a=ewxv(pitn)
 ENDIF
ENDDO !while
pitw2(t)=mw2-1
mt=pitt(t)+1
IF (mt.LE.pitn) THEN
 a=etxv(mt)
ELSE
 a=etxv(pitn)
ENDIF
DO WHILE ((mt.LE.pitn).AND.(ert(t).GT.a))
 mt=mt+1
 IF (mt.LE.pitn) THEN
  a=etxv(mt)
 ELSE
  a=etxv(pitn)
 ENDIF
ENDDO !while
pitt(t)=mt-1
erz(t)=dfloat(pitz(t))/pitn
erc(t)=dfloat(pitc(t))/pitn
erw(t)=dfloat(pitw(t))/pitn
erw2(t)=dfloat(pitw2(t))/pitn
ert(t)=dfloat(pitt(t))/pitn
err2=dfloat(t-skip)/dfloat(tt+1-skip) ! Equi-angular line
```

```
! Output
       WRITE(3,'(19F15.10)') y(t),rv(t),err2,erx(t),erv(t),ert(t),erw(t),
     &        erz(t),erc(t),erw2(t),hth(t),htm(t),htz(t),htx(t),htv(t),
     &        htt(t),htw(t),htc(t),htw2(t)
       ENDDO !t
! Day-Loop; End
! De-allocate arrays
      DEALLOCATE(ewxv,ecxv,ezxv,etxv)
      WRITE(6,*) 'Residual analysis output written to '//output
      END SUBROUTINE PIT_Output
```

BIBLIOGRAPHY

Abramowitz, S., and Stegun, I.A. (1965). Handbook of Mathematical Functions. Dover, New York.

Andersen, T.G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39, 885–905.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Ebens, H. (2001). The distribution of stock return volatility. *Journal of Financial Economics*, 61, 43-76.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, H. (2000). Exchange rate returns standardized by realized volatility are (nearly) Gaussian. *Multinational Finance Journal*, 4, 159-179.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, H. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96, 42-55.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, 71, 579-625.

Andersen, T.G., Chung, H. and Sørensen, B.E. (1999). Efficient method of moments estimation of a stochastic volatility model: a Monte Carlo study. *Journal of Econometrics*, 91, 61-87.

Andersson, J. (2001). On the normal inverse Gaussian stochastic volatility model. *Journal of Business and Economic Statistics*, 19, 44-54.

Andrews, D. and Mallows, C. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society*, Vol. 36, 99-.

Atkinson, A.C. (1982). The simulation of generalised Inverse Gaussian and Hyperbolic random variables. *Siam Journal of Scientific and Statistical Computing*, 3, 502-515.

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution, *Journal of the Royal Statistical Society*, Series B, Vol. 61, pp.579-602.

Azzalini A. and Capitanio, A. (2003). Distributions generated by permutation of symmetry with emphasis on a multivariate skew *t* distribution. *Journal of the Royal Statistical Society*, Series B, Vol. 65, 367-389.

Barndorff-Nielsen, O.E. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceeds of the Royal Society of London*, Series A, Vol. 353, 401-419.

Barndorff-Nielsen, O.E. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics*, 5, 151-157.

Barndorff-Nielsen, O.E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24, 1-13.

Barndorff-Nielsen, O.E., Blæsild, P. and Seshadri, V. (1992). Multivariate distributions with generalized inverse Gaussian marginals, and associated Poisson mixtures. *The Canadian Journal of Statistics*, Vol. 20 No. 2, 109-120.

Barndorff-Nielsen, O.E., Blæsid, P., Jensen, J.L. and Sorensen, M. (1985). The fascination of sand. In: A celebration of Statistics, Atkinson, A.C. and Feinberg, S.E. editors. Springer-Verlag, New York, 57-87.

Barndorff-Nielsen, O.E., and Prause, K. (2001). Apparent Scaling. *Finance and Stochastics*, 5, 103-113.

Bates, D.S. (1996). Jumps and stochastic volatility: exchange rate processes implicit in Deutsche Mark options. *The Review of Financial Studies*, 9 (1), 69-107.

Blæsild, P. (1978). The shape of the generalized inverse Gaussian and hyperbolic distributions. Research Report, 37, Aarhus, Denmark.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, Vol. 31(3), 307-327.

Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics*, Vol. 69, 542–547.

Bollerslev, T., Chou, R.Y., and Kroner, K.F. (1992). ARCH modelling in finance: a review of the theory and empirical evidence. *Journal of Econometrics*, 39, 5-59.

Bollerslev, T., Engle, R.F., and Nelson, D.B. (1994). ARCH models. In: Handbook of Econometrics, Vol. 4, Engle, R.F. and McFadden, D.C. editors. North-Holland, Amsterdam, 2959-3038.

Bollerslev, T. and Wooldridge, J.M. (1992). Quasi-maximum likelihood estimation of dynamic models with time varying covariances. *Econometric Reviews*, 11, 143-172.

Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control. Holden-Day, Oakland.

Chhikara, R.S. and Folks, J.L. (1974). Estimation of the Inverse Gaussian Distribution Function. *Journal of the American Statistical Association*, 69, 250-254.

Chhikara, R.S. and Folks, J.L. (1989). The Inverse Gaussian Distribution. Marcel-Dekker, Inc., New York.

Clark, P.K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41, 135-155.

Cox, D.R. and Miller, M.D. (1965). The theory of stochastic processes. Methuen & Co. Ltd., London.

Dagpunar, J.S. (1989). An easily implemented generalised Inverse Gaussian generator. Communications in Statististics: Simulation and Computation, 18(2), 703-710.

Danielsson, J. (1994). Stochastic volatility in asset prices: Estimation with simulated maximum likelihood. *Journal of Econometrics*, 64, 375-400.

Durbin, J. and Watson, G.S. (1951). Testing for Serial Correlation in Least Squares Regression II. *Biometrika*, 38, 159–178.

Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987-1008.

Engle, R.F. (2002). New Frontiers for ARCH. *Journal of Applied Econometrics*, 17, 425–446.

Engle, R.F. (2004). Risk and Volatility: Econometric models and financial practice. *American Economic Review*, 94 (3), 405-420.

Engle, R.F. and Ng, V.K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance,* 48(5), 1749-1778.

Engle, R.F. and Ishida, I. (2002). Forecasting Variance of Variance: The Square-Root, the Affine, and the CEV Garch Models. *Department of Finance Working Papers*, New York University.

Fama, E.F. (1963). Mandelbrot and The Stable Paretian Distribution. *Journal of Business*, 36, 420-429.

Fernándes, C. and Steel, M.F.J. (1998). On Bayesian Modeling of Fat Tails and Skewness. *Journal of the American Statistical Association*, 93, 359-371.

Fleming, J., Kirby, C. and Ostdiek, B. (2004). Stochastic volatility, trading volume, and the daily flow of information. *Journal of Business*, 79, 1551-1590.

Forsberg, L. (2002). On the Normal Inverse Gaussian Distribution in modeling volatility in the Financial Markets. Doctoral Thesis, Studia Statistica Upsaliensia 5, Christofferson, A. and Joreskog, K.G. editors. Uppsala University Library.

Forsberg, L. and Bollerslev, T. (2002). Bridging the gap between the distribution of realized (Ecu) volatility and ARCH modelling (of the Euro): The GARCH-NIG Model. *Journal of Applied Econometrics*, 17, 535-548.

Franses, P.H. and MacAleer, M. (2002). Financial Volatility: An Introduction. *Journal of Applied Econometrics*, 17, 419-424.

Frey, R. and McNeil, A.J. (2003). Dependent defaults in models of portfolio credit risk. *Journal of Risk*, 6(1), 59-92.

Ghysels, E., Harvey, A.C., and Renault, E. (1996). Stochastic Volatility. In: Handbook of Statistics, Vol. 14, Maddala, G.S. and Rao, C.R. editors. Elsevier Science B.V., Amsterdam.

Gigli, A. (2002). Market Volatility: What do We Know?. Institute of Finance, USI, Lugano (CH). October 16, 2002, 1.

Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, 40, 237–264.

Goodhart, C. A. E. and O'Hara, M. (1997). High frequency data in financial markets: Issues and applications. *Journal of Empirical Finance*, 4 (2), 73-114.

Gouriéroux, C. (1997). ARCH models and financial applications. Springer, New york.

Hansen, P.R. and Lunde, A. (2001). A comparison of volatility models: Does anything beat a GARCH(1,1)? *Working Paper Series No. 84,* CAF, University of Aarhus.

Heston, S.L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6 (2), 327-43.

Hida, T. (1980). Brownian Motion. Springer-Verlag.

Jacquier, E., Polson, N.G. and Rossi, P.E. (1194). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics*, 12, 371-389.

Jensen, M.B. and Lunde, A. (2001). The NIG-S&ARCH model: a fat-tailed, stochastic and autoregressive conditional heteroskedastic volatility model. *Econometrics Journal*, 4(2), 319-342.

Jorgensen, B. (1982). Statistical Properties of the Generalized Inverse Gaussian Distribution. *Lecture Notes in Statistics*, Vol. 9, Springer, Berlin.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). Continuous univariate distributions, Vol 2. 2nd Ed. J. Wiley and Sons, New York.

Jones, M.C. and Faddy, M.J. (2003). A skew extension of the t distribution, with applications. *Journal of the Royal Statistical Society*, Series B, Vol. 65, 159-74.

Kendall, M.G. and Stuart, A. (1969). The Advanced Theory of Statistics, Volume 1: Distribution Theory, 3rd Edition. Griffin, London.

Kinderman, A.J. and Monahan, J.F. (1977). Computer generation of random variables using the ratio of uniform deviates. *ACM Transactions on Mathematical software*, 3, 257-260.

Law, A.M. and Kelton, W.D. (2000). Simulation Modeling and Analysis, 3rd Edition. McGraw-Hill.

Lehman, E.L. and Casella, G. (1998). Theory of Point Estimation, 2nd Edition. Springer, New York.

Li, W.K., Ling, S. and McAleer, M. (2002). Recent theoretical results for time series models with GARCH errors. *Journal of Economics Surveys*, 16, 245-269.

Liesenfeld, R. (2001). A generalized bivariate mixture model for stock price volatility and trading volume. *Journal of Econometrics*, 104, 141-178.

Liesenfeld, R. and Richard, J.F. (2003). Univariate and multivariate stochastic volatility models: Estimation and diagnostics. *Journal of Empirical Finance*, 10, 505-531.

Lildholdt, P.M. (2002). Estimation of GARCH models based on open, close, high, and low prices. *Working Paper 128*, Centre for Analytical Finance, Aarhus School of Business.

Lillestøl, J. (2000). Risk analysis and the NIG distribution. *The Journal of Risk*, 4(2), 41-56.

Mandelbrot, B.B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, 36, 394-416.

McNeil, A.J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroskedastic financial time series: an extreme value approach . *Journal of Empirical Finance*, 7, 271-300.

McNeil, A.J., Frey, R. and Embrechts, P. (2005). Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, New Jersey.

Melino, A. and Turnbull, S.M. (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, 45, 239-265.

Michael, J.R., Schucany, W.R. and Haas R.W. (1976). Generating random variates using transformations with multiple roots, *The American Statistician*, 30, 88-89.

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

Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *Journal of Business*, Vol. 53(1), University of Florida, 61-65.

Prause, K. (1999). The generalized hyperbolic model: Estimation, financial derivatives, and risk measures. Dissertation, University of Freiburg.

Rice, J.A. (1995). Mathematical Statistics and Data Analysis. Duxbury Press, Belmont, California.

Rogers, L.C.G. and Satchell, S.E. (1991). Estimating Variance from High, Low and Closing Prices. *Annals of Applied Probability*, 1, 504-512.

Rydberg, T.H. (1997). The normal inverse Gaussian Levy process: simulation and approximation. *Communications in Statistics: Stochastic Models*, 13, 887-910.

Shuster, J. (1968). On the inverse Gaussian distribution function. *Journal of the American Statistical Association*, 63, 1514-1516.

Sichel, H.S. (1974). On a distribution representing sentence-length in written prose. *Journal of the Royal Statistical Society*, Series A, Vol. 137(1), 25–34.

Sichel, H.S. (1975). On a distribution law for word frequencies. *Journal of the American Statistical Society*, 70(351), 542–547.

Straumann, D. (2005). Estimation in conditionally heteroskedastic time series models. Lecture notes in Statistics, Volume 181. Springer, Berlin.

Taylor, S. and Xu, X. (1997). The Incremental Volatility Information in One Million Foreign Exchange Quotations. *Journal of Empirical Finance*, 4, 317–340.

Venter, J.H., and de Jongh, P.J. (2002). Risk estimation using the normal Inverse Gaussian distribution. *The Journal of Risk*, 4(2), 1-23.

Venter, J.H. and de Jongh, P.J. (2004). Selecting an innovation distribution for GARCH models to improve efficiency of risk and volatility estimation. *The Journal of Risk*, 6(3), 27-53.

Venter, J.H., de Jongh, P.J. and Griebenow, G. (2004). The BIG-GARCH model for intra-day returns. North-West University, South Africa, Manuscript FABWI-N-WST:2004-125.

Venter, J.H., de Jongh, P.J. and Griebenow, G. (2005). NIG-GARCH models based on open, close, high and low prices. *South African Statistical Journal*, 39, 173-195.

Xu, X.E., Chen, P. and Wu, C. (2005). Time and dynamic volume-volatility relation. *Journal of Banking and Finance* Forthcoming.