Ensayos Revista de Economía–Volumen XXX, No. 2, noviembre 2011, pp. 1-28

Modelling the Density of Inflation Using Autoregressive Conditional Heteroscedasticity, Skewness, and Kurtosis Models

Doaa Akl Ahmed*

Fecha de recepción: 21 XI 2010 Fecha de aceptación: 28 III 2011

Abstract

The paper aimed at modelling the density of inflation based on time-varying conditional variance, skewness and kurtosis model developed by Leon, Rubio, and Serna (2005) who model higher-order moments as GARCH-type processes by applying a Gram-Charlier series expansion of the normal density function. Additionally, it extended their work by allowing both conditional skewness and kurtosis to have an asymmetry term. The results revealed the significant persistence in conditional variance, skewness and kurtosis which indicate high asymmetry of inflation. Additionally, diagnostic tests reveal that models with nonconstant volatility, skewness and kurtosis are superior to models that keep them invariant.

Keywords:. inflation targeting, conditional volatility, skewness and kurtosis, modelling uncertainty of inflation.
JEL Classification: C13, E31, E37.

Resumen

El objetivo de este artículo es modelar la densidad de la inflación con base en la variación temporal de la varianza condicional, asimetría y curtosis del modelo desarrollado por León, Rubio y Serna (2005), en donde se aplicó la serie de expansión Gram-Charlier de la función de densidad normal para modelar procesos tipo GARCH con momentos de mayor orden. Además, este artículo proporciona una extensión al artículo mencionado, ya que permite que tanto la asimetría y la curtosis condicionales incluyan un término de asimetría. Los resultados revelan persistencia significativa en la varianza condicional, asimetría y curtosis, lo cual sugiere alta asimetría en la

* University of Leicester, University Road, Leicester, LE1 7RH, United Kingdom, and University of Benha, Egypt.
Email: da88@le.ac.uk

inflación. Por su parte, las pruebas de diagnóstico revelan que los modelos con volatilidad no constante, asimetría y curtosis son superiores a los modelos que las asumen invariables.

Palabras Clave: inflación objetivo, volatilidad condicional, asimetría y curtosis, modelación de la inflación en condiciones de incertidumbre.
Clasificación JEL: C13, E31, E37.

Introduction

Exploring the relation between inflation and its higher-order moments is quite important for central banks especially under an Inflation Targeting (IT) framework. That is because policymakers are increasingly worrying about complete density forecasts that allow a much richer setting of anti-inflation policy. Therefore, the paper aims at exploring the relation between CPI inflation and its higher-order moments that allows better understanding of the risks involved in inflation.

As indicated by Engle (1982), unpredictability of inflation causes high levels of welfare loss associated with inflation. The reason behind that is however, the costly expected inflation due to institutional rigidities, government intervention, and transaction costs. On the one hand, these costs will be minimized in the long-run if agents use different forms of indexation. On the other hand, the lack of ability to predict future inflation affects risk averse agents as it damages the efficiency of allocation decisions between current and future periods, due to uncertainty; even though prices and quantities are perfectly flexible in all markets. Additionally, as many countries have adopted an inflation targeting (IT) regime, this requires the existence of efficient inflation forecasts models. Given that the Central Bank of Egypt (CBE) announced its intention to adopt an IT framework to anchor its monetary policy when the basic prerequisites are satisfied (CBE, 2005), it must have accurate models to forecast future inflation.

Literature on inflation forecasting is still very limited in Egypt. However, some studies analysed and estimated the sources of inflation but it did not assess the ability of these models to forecast inflation. Noureldin (2005) assessed the robustness of three alternative approaches to forecast inflation in Egypt. These three approaches are output gap (Phillips curve), money gap model, and Vector Autoregressive (VAR) model. However, point forecast does not provide a full description of the uncertainty associated with the forecast. Actually, the Business and Economic Statistics Section of the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER) initially published the first series of density forecasts in macroeconomics in 1968. They jointly started a quarterly survey of macroeconomic forecasters in the United States known as the ASA-NBER survey (Tay and Wallis, 2000). Additionally, the Bank of England published a density forecast of inflation in its quarterly Inflation Report since February 1996 (Wallis, 2004).

An examination of the monthly Egyptian CPI inflation data reveals that it exhibits high degree of volatility, which is a necessary condition to apply models that allow for time-varying conditional variance, skewness, and kurtosis. Forecasting monthly inflation is highly needed by central banks as many central banks including CBE publish monthly bulletin including inflation data. For inflation targeters, it is more appropriate to publish monthly forecasts to enable monetary authority to better decision making in terms of altering instruments. To investigate the importance of including second, third, and fourth moments, six models have been estimated assuming different distributions for the error term. The first two models are GARCH-M, and its threshold (TARCH-M) extension assuming a normal distribution for the error term while the third is a GARCH-M model when the error term follows a GED distribution.

Nevertheless, the main contribution of the current paper is modelling the relationship between inflation and its second, third, and fourth moments that represent the full risks involved in inflation. Thus, the other models employed here are developed versions of GARCH and TARCH models that allow for jointly modelling the relation between inflation, its volatility, skewness, and kurtosis. This development is attributed to Leon, Rubio and Serna (2005), who model time-varying second, third and fourth moments jointly by applying a Gram-Charlier (G-C) series expansion of the normal density function. They extended the basic GARCH and NGARCH models to employ their methodology to different stock indices and exchange rates. Consequently, I will apply their methodology using GARCH-M and TARCH-M specification in measuring and forecasting inflation in the case of Egypt.

However, this paper differs from those of Leon *et al.* (2005) in three points. First, it uses inflation instead of stock returns and exchange rates. Second, current paper applies a TARCH specification of the conditional variance equation instead of NGARCH applied by them. Moreover, I allow the conditional skewness and kurtosis equation to have an asymmetry term (*i.e.*, TARCH specification for conditional third and fourth moments). The results show that specifications that allow skewness and kurtosis to vary with time outperform specifications that keep them constant. However, allowing both conditional skewness and kurtosis to follow a TARCH structure is inferior to allowing them to have a GARCH-type process. The paper is structured as follows. Section 1 is devoted to review the existing literature. Section 2 presents the different models while the preliminary check for the data, analysis of the results and comparison of different models is the core of section 3. Finally, section 4 applies the methodology to Mexico inflation data to show the applicability of the model to other economies. Finally, concludes and draws some policy implications.

1. Literature Review

As indicated before modelling the relation between inflation and its higher-order moments is quite important for policymakers to provide better understanding of the uncertainty of inflation. Regarding the relation between inflation and its volatility, Friedman (1977) asserts that high inflation leads to inflation that is more variable. This inflation uncertainty is costly since it distorts relative prices and increases risk in nominal contracts (Berument, Metin-Ozcan, and Neyapti, 2001). From the empirical viewpoint, Engle (1982) finds that for some kinds of data including inflation, the variance of the disturbance term is not stable as usually assumed by OLS model. Assuming that the error term exhibits heteroscedasticity, however with a zero and serially uncorrelated mean, he found that UK inflation follows Autoregressive Conditional Heteroscedasticity (ARCH) process (*i.e.*, variances conditional on past value of the error term are time-varying while unconditional variance is constant). He uses Maximum Likelihood (ML) methodology to estimate this model and claimed that the ML estimator is more efficient than the OLS estimator. To detect the existence of ARCH effects, he applied a Lagrange Multiplier (LM) test based on the autocorrelation of the squared OLS residuals.

However, the basic ARCH model has some shortcomings including the absence of clear approach to choose the suitable number of lags of the squared residuals to be included in the model. Additionally, this number of lags may be quite large leading to non-parsimonious model and violation of the nonnegativity constraints on variance parameters. Moreover, it assumes that the current conditional volatility depends only on the past values of residuals squared which may be unrealistic assumption as volatility response to positive and negative shocks are not similar (Engle, 1995; Rachev *et al*, 2007; and Brooks, 2002).

The basic ARCH model has been extended many times. Bollerslev (1986) introduced generalized version of the ARCH process by allowing the conditional variance to be an Autoregressive Moving Average (ARMA) process which permits a more flexible lag structure without the violation of the nonnegativity restrictions.

However, the basic GARCH model is criticised as it assumes that the response of variance to negative and positive shocks is similar. Due to the observed asymmetry of the variance response to shocks with different signs, some variant models have been developed to account for this asymmetry. The first variant is the exponential GARCH (EGARCH) model suggested Nelson (1991) to permit conditional volatility to be a function of both the size and the sign of lagged residuals assuming that the residuals follow Generalized Error Distribution (GED). However, this distribution allows shocks of different signs to have different impact on volatility, but it is still symmetric like the normal distribution (Harvey and Siddique, 1999). Additionally, Glosten, Jagannathan and Runkle (1993) introduced a formula that captures the leverage effect of financial time series, namely threshold ARCH (TARCH) or GJR specification[1].

Although ARCH family models are quite useful in modelling time-series variation in conditional volatility, these models assume that the conditional distribution are time-varying only in the first two moments and ignore the information content in higher-order moments (Chaudhuri, Kim, and Shin, 2011). To fill this gap, Harvey and Siddique (1999) developed a new approach to estimate nonconstant conditional skewness. They extended the traditional GARCH (1,1) model by explicitly modelling the conditional variance and skewness using maximum likelihood framework assuming that the standardised errors follow noncentral t-distribution. To allow for nonconstant conditional kurtosis, Leon *et al.* (2005) developed the methodology of Harvey and Siddique (1999) by jointly modelling time-varying variance, skewness and kurtosis (GARCHSK model) assuming that the error term is derived by G-C series expansion of the normal density function which is easier to estimate than the noncentral t-distribution suggested by Harvey and Siddique (1999).

Chaudhuri, Kim, and Shin (2011) introduce a semi-parametric Functional Autoregressive (FAR) model for forecasting a time-varying distribution of the sectoral inflation rates in the UK. Their approach employs the autoregressive operator to specify the time-dependence of the distribution function and thus allows all the moments to depend on all the past moments. Thus, they do not impose particular moment specifications like those in the conventional parametric models.

Concerning the relationship between inflation and skewness, it could be investigated using two models that differ in the degree of flexibility in the

[1] For more details about the different extensions of ARCH/GARCH models, see Bollerslev (2008).

economy. Under a sticky price model, Ball and Mankiw (1995) argue that when the economy is exposed to a supply shock, firms can adjust their prices but they have to face a menu cost. Thus, the firm will change its price if the shock is large enough and the menu cost is less than the loss resulting from keeping the price unchanged. Therefore, large shocks have unequal effects on the price level. They conclude that aggregate inflation depends on the distribution of relative price changes. As the distribution of the relative price shock is asymmetric, it will cause the distribution of aggregate inflation to be asymmetric as well. This asymmetric distribution for the relative price shocks cause temporary fluctuations in the mean of prices and hence will lead to a positive correlation between the mean and skewness of the price-change distribution. However, the model assumes that the mean-skewness correlation vanishes in the long-term since this correlation is attributed to short-run considerations. Therefore, the correlation must be declined and die out as the time length used to measure price changes increases.

On the other hand, under a flexible price model, Balke and Wynne (1996) show that allowing the prices to be flexible does not capture the observed relation between mean and skewness of the relative price changes. However, when the model is amended to include the input-output relationship between sectors as well the variance of productivity (or supply) shocks, the mean-skewness relation is captured by the flexible price model. Thus, a supply shock could affect prices through two channels. The first one, it changes the level of output and hence the aggregate price level given a certain level of money supply. The other channel works through the inter-firm purchases that causes different price changes on different sectors of the economy. The relative price changes are conditional on the influence of supply shock on the level of productivity in a certain sector. Since the input-output structure is asymmetric, the distribution of price changes will be skewed. Thus, this model assumes a positive correlation between the first and the third moment of inflation. Moreover, they assume that this relation should persist or even it may be strengthened in the long-run. Therefore, the skewness of inflation may be of a great importance in investigating and forecasting future inflation.

In fact, Bryan and Cecchetti (1996) reported fat-tailed properties in inflation data. Additionally, Roger (2000) found evidence towards right skewness. In addition, Chaudhuri, Kim, and Shin (2011) found that the mean inflation is positively correlated with variance and skewness. These results suggest that a greater attention must be paid to increases than decreases of inflation rates.

Given the growing importance of accounting for nonconstant higher order moments, I will apply the specification proposed by Leon *et al.* (2005) for conditional third and fourth moments to Egyptian inflation data. This will be followed by examining the possibility to apply the methodology to Mexican CPI inflation.

2. Empirical Models

This section presents the basic GARCH model briefly as well as the TARCH extension to account for the leverage effect. Additionally, a GARCH-M specification will be presented in short. These standard models are presented as their parameter estimates will be used in the developed models that allow conditional skewness and kurtosis to vary across time. Then, these extended models of Leon *et al.* (2005) will be introduced in details.

2.1 Basic Models of Time-Varying Conditional Volatility

As indicated in the introduction, Bollerslev (1986) extended the basic ARCH model to relate the conditional variance to both past squared errors and past conditional variances. The GARCH(1,1) model has the following specification of the conditional variance

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} \qquad \text{(i)}$$

Where h_t is the conditional variance, h_{t-1} is the past volatility which is used as a measure of variance persistence and $\varepsilon_{t-1}^2$is the past squared errors.

In order to ensure that the conditional variance is strictly positive, the following inequality restrictions are to be imposed: $\beta_0 \geq 0$, $\beta_1 \geq 0$, $\beta_2 \geq 0$. Additionally, to insure stationarity, it is also required that $\beta_1 + \beta_2 < 1$ where the persistence of variance becomes higher as β_2 approaches 1.

One of the key restrictions of GARCH(p,q) models is that they enforce a symmetric response of volatility to positive and negative shocks. GJR specification that captures the leverage effect of financial time series could be written as

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} + \beta_3 \varepsilon_{t-1}^2 (\varepsilon_{t-1} < 0) \qquad \text{(ii)}$$

According to the TARCH model, the conditions $\beta_0 > 0$, $\beta_1 > 0$, $\beta_1 + \beta_3 > 0$, $\beta_2 \geq 0$ are sufficient to ensure a strictly positive conditional variance. The asymmetry parameter β_3 is allowed to be of either sign to capture the asymmetric effects. This parameter measures the contributions of shocks to both short run persistence $(\beta_1 + \beta_3/2)$ and long run persistence $(\beta_1 + \beta_2 + \beta_3/2)$. Another interpretation of the relation between the mean inflation and its uncertainty allows the conditional variance to be a regressor in the mean equation. This GARCH in mean specification denoted GARCH-M add another term in the equation of the mean as follows:

$$\pi_t = \mu h_t + \sum_{i=1}^{n} \alpha_i \pi_{t-i} + \varepsilon_t \quad \text{(iii)}$$

Where π_t refers to inflation, h_t is the conditional volatility. Actually, the relation between inflation, volatility and price dispersion has been investigated using GARCH-M specification (Grier and Perry, 1996). Their results suggest that inflation volatility is superior to trend inflation in investigating price dispersion. Additionally, Wilson (2006) employs an EGARCH-M model to explain the relation between inflation, its volatility and output gap. Their results suggested that higher uncertainty do raise inflation and reduce output, which supports Friedman (1977) argument.

2.2. Modelling Conditional Variance, Skewness and Kurtosis[2]

Leon *et al.* (2005) developed a new approach allowing for modelling time-varying variance, skewness and kurtosis jointly as a GARCH process. The employed likelihood function, based on the series expansion of the normal density function is less complicated to estimate in comparison with the likelihood function proposed by Harvey and Siddique (1999) that assumes non-central t distribution for the model errors.

First, an inflation model is specified as GARCH(1,1)-M or TARCH (1,1)-M. Then, it is included a GARCH(1,1) specification for both conditional nonconstant skewness and kurtosis. Let GARCHSK-M refers to the model when the conditional variance is derived by a GARCH specification while TARCHSK-M when conditional variance is derived by the TARCH (1,1)-M model. In addition, denote the specification that allows for an asymmetry term in the skewness and kurtosis equation by TARCHTSK-M. Thus, the different models are specified as follows

Mean equation:

$$\pi_t = \mu h_t + \sum_{i=1}^{n} \alpha_i \pi_{t-i} + \varepsilon_t \; \varepsilon_t \approx (0, \sigma_\varepsilon^2) \quad (1.1)$$

$$\varepsilon_t = \eta_t \sqrt{h_t} \; ; \eta_t \approx (0,1) \qquad E(\varepsilon_t | I_{t-1}) \approx (0, h_t)$$

---

[2] This section is mainly based on Leon *et al.* (2005) and their development to the GARCH-type model of skewness and kurtosis.

Variance (GARCH):

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} \tag{1.2}$$

Variance (TARCH):

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} + \beta_3 \varepsilon_{t-1}^2 (\varepsilon_{t-1} < 0) \tag{1.3}$$

Skewness (GARCH):

$$s_t = \gamma_0 + \gamma_1 \eta_{t-1}^3 + \gamma_2 s_{t-1} \tag{1.4}$$

Skewness (TARCH):

$$s_t = \gamma_0 + \gamma_1 \eta_{t-1}^3 + \gamma_2 s_{t-1} + \gamma_3 \eta_{t-1}^3 (\eta_{t-1} < 0) \tag{1.5}$$

Kurtosis (GARCH):

$$k_t = \delta_0 + \delta_1 \eta_{t-1}^4 + \delta_2 k_{t-1} \tag{1.6}$$

Kurtosis (TARCH):

$$k_t = \delta_0 + \delta_1 \eta_{t-1}^4 + \delta_2 k_{t-1} + \delta_3 \eta_{t-1}^4 (\eta_{t-1} < 0) \tag{1.7}$$

Where ε_t is the error term, η_t is the standardised residuals, h_t, s_t, and k_t are conditional volatility, skewness and kurtosis corresponding to η_t respectively. They establish that $E_{t-1}(\eta_t) = 0, E_{t-1}(\eta_t^2) = 1, E_{t-1}(\eta_t^3) = s_t$ and $E_{t-1}(\eta_t^4) = k_t$. First, two basic models are estimated, a GARCH (1,1)-M (equations (1.1) and (1.2)) and a TARCH (1,1)-M (equations (1.1) and (1.3)). This followed by models with nonconstant higher order moments, GARCHSK-M (equations (1.1), (1.2), (1.4) and (1.6)), TARCHSK-M (equations (1.1), (1.3), (1.4) and (1.6)), and a TARCHTSK-M (equations (1.1), (1.3), (1.5) and (1.7)).

They employed G-C series expansion of the normal density function and truncated at the fourth moment to get the following density function for the standardised errors

$$\begin{aligned} f(\eta_t | I_{t-1}) &= \phi(\eta_t) \left[1 + \frac{s_t}{3!}(\eta_t^3 - 3\eta_3) + \frac{k_t - 3}{4!}(\eta_t^4 - 3\eta_t^2 + 3) \right] \\ &= \phi(\eta_t)\psi(\eta_t) \end{aligned} \tag{2}$$

Where $\phi(.)$ denotes the Probability Density Function (pdf) corresponding to the standard normal distribution. Since some parameter estimates in equation (1) may lead to negative value of $f(.)$ due to the component $\psi(.)$, therefore, $f(.)$ is not a real density function. Additionally, the integral of $f(.)$ on R is not equal to one. Therefore, Leon *et al.* (2005) introduce a true pdf, by squaring the polynomial part $\psi(.)$, and dividing by the integral of $f(.)$ over

R to assure that the density integrates to one. The resulting form of pdf is as follows:

$$w(\eta_t|I_{t-1}) = \phi(\eta_t)\psi^2(\eta_t)/\Gamma_t \quad (3)$$

Where

$$\Gamma_t = 1 + \frac{s_t^2}{3!} + \frac{(k_t - 3)^2}{4!}$$

Thus, the logarithm of the likelihood function for one observation corresponding to the conditional distribution $\varepsilon_t = \eta_t\sqrt{h_t}$, whose pdf is $\sqrt{h_t}w(\eta_t|I_{t-1})$ could be reached after deleting the redundant constants as follows

$$l_t = -\frac{1}{2}lnh_t - \frac{1}{2}\eta_t^2 + ln(\psi^2(\eta_t)) - ln(\Gamma_t) \quad (4)$$

One advantage of this likelihood function is the similarity with the standard normal density function in addition to two adjustment terms to account for time-varying third and fourth moments. What is more, the aforementioned developed density function in equation (3) nests the normal density function (when s_t= zero and k_t= 3). Thus, the restrictions imposed by the normal density functions (*i.e.,* $\gamma_0 = \gamma_1 = \gamma_2 = \delta_0 = \delta_1 = \delta_2 = 0$) could be tested by conducting a likelihood ratio test.

3. Empirical Results

3.1. Data and Preliminary Check

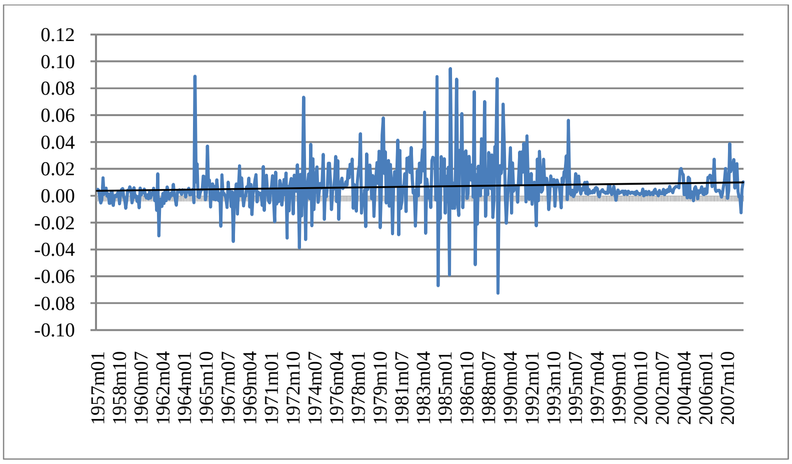
The data of monthly CPI is collected from IMF's International Financial Statistics (IFS) and cover the period 1957:1 to 2009:2. Inflation data is computed as monthly changes in the logarithm of CPI. The sample is chosen to include the largest number of available observations to provide more accurate results. Table 1 gives the basic descriptive statistics for the data. It is clearly shown that the distribution of the data is right-skewed and leptokurtic (*i.e.*, the data are not normally distrusted as indicated by Jarque-Bera test statistic).

Table 1
Descriptive statistics of monthly inflation rate for the Period (1957:1 to 2009:2)

Mean	0.0068
Median	0.0039
Maximum	0.0946
Minimum	-0.0725
Std. Dev.	0.0166
Skewness	1.0967
Kurtosis	10.2498
Jarque-Bera	1494.0770
Probability	0.0000

Source: Own calculations.

Figure 1
Monthly inflation rate for the period (1957:1 to 2009:2)



Source: Author's calculation based on CPI data from IFS.

It will be examined the dynamics structure in the conditional mean before estimating models concerned with the dynamics of conditional volatility, skewness, and kurtosis. To explore the dynamics of the conditional mean, the correlogram of inflation have been analysed as guidance for selecting the appropriate mean specification.

According to Brooks (2002) a given autocorrelation coefficient is classified as significant if it is outside a range of $\pm 1.96 \times 1/\sqrt{N}$, where N is the number of observations. In this case, it would imply that a correlation coefficient is classified as significant if it were outside the band of -0.08 and

0.08. Examining the correlogram of the data reveals that autocorrelation and partial autocorrelation coefficients are significant under that rule for first, second, sixth, ninth and twelfth lags. Therefore, an ARMA process seems to be suitable, although it is hard to determine the appropriate order precisely given these results. Thus, the information criteria are employed to determine the appropriate order.

Accordingly, different specification for the mean equation has been applied using different orders of AR and MA terms. In the underlying case, criteria choose different models. That is while AIC selects an ARMA (3,3) specification of the mean equation, SIC chooses ARMA(2,1). However, by checking LJUNG-BOX Q-statistic to see if the models are free from autocorrelation leads to the rejection of the null hypothesis of no autocorrelation. Therefore, both specifications are not valid. Re-examining the values of both criteria show that many different models provide almost identical values of the information criteria, which indicates that the chosen models do not provide particularly sharp characteristics of the data and other specifications could fit the data almost as well.

As a result, I have estimated many different specifications using different significant lags detected from the correlogram. The results show that any specification that does not include second and twelfth lags would exhibit serial correlation between the residuals. According to SIC, the chosen model should include first, second and twelfth lags of inflation. Diagnostic checks reveal the absence of serial autocorrelation amongst the residuals while it exists in the sequences of squared, cubed, and the fourth power of residuals. Additionally, ARCH LM test which indicates the existence of ARCH effects in the residuals. Therefore, a model that allow second, third, and fourth order moments to be time-varying would be more appropriate in modelling inflation.

As the likelihood function is highly nonlinear, good starting values of the parameters are required. Thus, the models should be estimated in steps starting from simpler models that nested in the complicated ones. In other words, the estimated parameters of the simpler models are used as starting values for more complex ones. Accordingly, I start by modelling inflation using basic GARCH(1,1)-M model followed by TARCH(1,1)-M model to test the asymmetry of volatility response to the sign of the shock to inflation. It is worth noting that the variance equation is allowed to include two dummies, *d74* and *July*. The first dummy captures the effects of shifting to the open door policy in 1974 that leads to a high increase in the inflation rate. The second dummy is included to capture the beginning of the financial year that witnesses the annual increase in wages. Addition of these dummies to the volatility equation allows for exploring their effect on the variability of inflation. Furthermore, both dummies are essential to insure covariance stationarity in TARCH-M model and its extensions. Moreover, I have estimated a GARCH-M model with GED distribution for the error term. This is done to compare the effect of choosing a non-normal distribution of the error term with constant skewness and kurtosis with models that allow skewness and kurtosis to vary with time.

3.2. Results

Table 2 reports the results of two basic models, GARCH-M with normal distribution, GED distribution, and the GARCHSK-M model with time-varying conditional third and fourth moments. Results indicate a significant presence of conditional variance persistence as the parameter of lagged volatility is positive and significant across the different models. Thus, high conditional volatility leads to higher conditional volatility next periods. Additionally, the coefficient of volatility persistence decreases by allowing nonconstant conditional skewness and kurtosis in GARCHSK-M specification. Thus, a time-varying conditional third and fourth moments lowers the magnitude of both volatility persistence and of shocks to conditional variance. Moreover, allowing the error term to follow a GED distribution leads to the highest volatility persistence. Concerning the conditional skewness, it is found that skewness persistence is positive and significant however, its magnitude is much lower than the variance case. In addition, shocks to skewness are significant and less than shocks to conditional variance. Similarly, the conditional kurtosis equation indicates that months with high kurtosis are followed by months with high kurtosis as conducted from the positivity and significance of lagged kurtosis parameter. Moreover, the coefficient of lagged kurtosis is higher than that of the lagged skewness and it is close the variance persistence parameter in the basic GARCH-M model. Finally, shocks effect to kurtosis are the highest related to the effects of shocks to volatility and skewness.

With respect to dummies effect in the variance equation, *d74* is positive and significant in all cases. Additionally, *July* is negative in all models but insignificant in GARCH-M where the error term follows a GED distribution.

Results of models that allow for asymmetries are displayed in Table 3. First, the asymmetry parameter in the volatility equation β_3 is found to be positive however insignificant in the basic TARCH model. Allowing the conditional skewness and kurtosis to follow a GARCH-type process lowers β_3 and keeps it insignificant while it turns to be significant with a little negative magnitude when the conditional skewness and kurtosis follow a TARCH-type process. Therefore, an unexpected decline in inflation has higher effect on volatility than that of unexpected rise according to the TARCHTSK model. On the other hand, both negative and positive shocks have the same effect on inflation variability under the basic TARCH-M and TARCHSK-M model where skewness and kurtosis is derived by GARCH process.

Secondly, the shocks to inflation β_1, it is found to be significant in the standard TARCH-M model and allowing for time-varying third and fourth moments raises its magnitude while still significant. Additionally, the persistence parameter in the variance equation is significant in all models with the highest magnitude in TARCHTSK-M model where both skewness and kurtosis are allowed to have an asymmetry parameter. Regarding skewness equation, the persistence parameter is positive and significant although it is less than that of the variance equation. As before, shocks to conditional skewness are lower in magnitude relative to shocks to conditional variance. Furthermore, as the coefficient for lagged kurtosis is positive and significant, months with high kurtosis are followed by months with high kurtosis. In addition, shocks to kurtosis have small magnitude close to the magnitude of shocks to conditional skewness. Finally, the asymmetries parameters γ_3 and δ_3 are significant with tiny magnitudes and positive in the case of skewness; however, it is negative in the kurtosis equation.

In addition, the parameter of GARCH in mean is significant in all models. Moreover, this estimate declines when moving from the simpler to the advanced models. In other words, inclusion of time-varying conditional skewness and kurtosis leads to a decline in μ.

Concerning the specification of the models, the Ljung-Box Q-statistics for the both standardised residuals and its squares are insignificant for lag length even larger than 20. Thus, there is evidence that both the level and squares of standardised residuals do not exhibit any serial correlation. Furthermore, ARCH LM tests indicate the absence of any further ARCH effects in the standardised residuals.

To choose the best model, SIC criterion is set to be equal to $\ln(LML) - (q/2)\ln(N)$, where q is the number of estimated parameters, N is the number of observations, and *LML* is the value of the log likelihood function using the q estimated parameters. Then, the best model is the one with the highest SIC. According to SIC criterion, the specification that allow for nonconstant third and fourth moments without an asymmetric term while the variance is structured as GARCH process (TARCHSK-M) is the best model.

To sum up, these results support Friedman (1977) hypothesis concerning the positive correlation between inflation and its uncertainty, as volatility persistence and GARCH in mean coefficients are significant in all models.

Additionally, the results show the evidence of positive skewness that is consistent with Balke and Wynne (1996) that the mean-skewness correlation could persist even in the long-run.

Table 2

GARCH-M, GARCH-M (GED) and GARCHSK-M Models for inflation

Mean equation: $\pi_t = \mu h_t + \alpha_1 \pi_{t-1} + \alpha_2 \pi_{t-2} + \alpha_3 \pi_{t-12} + \varepsilon_t$

Variance equation: $h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} + \kappa_1 d74 + \kappa_2 July$

Skewness Equation: $s_t = \gamma_0 + \gamma_1 \eta_{t-1}^3 + \gamma_2 s_{t-1}$

Kurtosis Equation: $k_t = \delta_0 + \delta_1 \eta_{t-1}^4 + \delta_2 k_{t-1}$

Model		**GARCH-M**		**GARCH-M (GED)**		**GARCHSK-M**	
		estimate	**p-value**	**estimate**	**p-value**	**estimate**	**p-value**
Mean Equation	μ	16.57217	**0.0000**	13.94960	**0.0000**	14.65763	**0.0000**
	α_1	-0.076226	0.1434	0.075078	**0.0020**	-0.053183	**0.0000**
	α_2	0.139471	**0.0039**	0.102082	**0.0001**	0.149367	**0.0000**
	α_3	0.244509	**0.0000**	0.213028	**0.0000**	0.174258	**0.0000**
Variance equation	β_0	1.71×10^{-5}	**0.0000**	3.36×10^{-6}	**0.0740**	0.000494	**0.0000**
	β_1	0.135005	**0.0047**	0.061078	**0.0002**	0.107057	**0.0000**
	β_2	0.800171	**0.0000**	0.911308	**0.0000**	0.757790	**0.0000**
	k_1	2.48×10^{-5}	**0.1027**	1.87×10^{-5}	**0.0225**	-3.86×10^{-5}	**0.0000**
	k_2	-7.56×10^{-5}	**0.0000**	-1.69×10^{-5}	0.3909	-0.000253	**0.0000**
	GED			0.846461	**0.0000**		
Skewness Equation	γ_0					-0.270272	**0.0000**
	γ_1					0.076670	**0.0000**
	γ_2					0.640609	**0.0000**
Kurtosis Equation	δ_0					0.117602	**0.0000**
	δ_1					0.147469	**0.0000**
	δ_2					0.807013	**0.0000**
Log-likelihood		1684.490		1787.573		2542.654	
SIC		1655.879		1755.784		2494.969	
Ljung-Box Q-stat.							
Residuals (lag 20)		16.591	(0.679)	22.055	(0.338)	21.690	(0.358)
squared (lag 20)		2.1088	(1.000)	0.6610	(1.000)	26.019	(0.165)

The basic GARCH-M with normal distribution is estimated using Quasi Maximum Likelihood (Bollerslev and Wooldridge, 1992) while GARCH-M (GED) and GARCHSK-M models are estimated using ML estimation. All models are estimated using Marquardt algorithm. Significant p-values are indicated by bold.

Source: Own calculations.

Table 3

TARCH-M, TARCHSK-M, and TARCHTSK-M Models

Mean equation:	$\pi_t = \mu h_t + \alpha_1 \pi_{t-1} + \alpha_2 \pi_{t-2} + \alpha_3 \pi_{t-12} + \varepsilon_t$
Variance equation:	$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} + \beta_3 \varepsilon_{t-1}^2(\varepsilon_{t-1} < 0) + \kappa_1 d74 + \kappa_2 July$
Skewness (GARCH):	$s_t = \gamma_0 + \gamma_1 \eta_{t-1}^3 + \gamma_2 s_{t-1}$
Skewness (TARCH):	$s_t = \gamma_0 + \gamma_1 \eta_{t-1}^3 + \gamma_2 s_{t-1} + \gamma_3 \eta_{t-1}^3(\eta_{t-1} < 0)$
Kurtosis (GARCH):	$k_t = \delta_0 + \delta_1 \eta_{t-1}^4 + \delta_2 k_{t-1}$
Kurtosis (TARCH):	$k_t = \delta_0 + \delta_1 \eta_{t-1}^4 + \delta_2 k_{t-1} + \delta_3 \eta_{t-1}^4(\eta_{t-1} < 0)$

Model		**TARCH-M**		**TARCHSK-M**		**TARCHTSK-M**	
		estimate	**p-value**	**estimate**	**p-value**	**estimate**	**p-value**
Mean equation	μ	15.20880	0.0000	14.33151	**0.0000**	10.99486	**0.0000**
	α_1	-0.066469	0.2109	-0.027512	**0.0000**	-0.040836	**0.0000**
	α_2	0.149982	0.0015	0.164858	**0.0000**	0.080036	**0.0000**
	α_3	0.251071	0.0000	0.254065	**0.0000**	0.230157	**0.0000**
Variance equation	β_0	1.71×10^{-5}	0.0000	4.74×10^{-5}	**0.0000**	2.01×10^{-5}	**0.0000**
	β_1	0.076294	0.1155	0.135653	**0.0000**	0.120574	**0.0000**
	β_2	0.806646	0.0000	0.798135	**0.0000**	0.874757	**0.0000**
	β_3	0.113107	0.2040	0.001450	0.6793	-0.008974	**0.0113**
	k_1	2.59×10^{-5}	0.0477	0.000135	**0.0000**	3.28×10^{-5}	**0.0000**
	k_2	-7.73×10^{-5}	0.0000	-0.000184	**0.0000**	-5.45×10^{-5}	**0.0000**
Skewness equation	γ_0			-0.046571	**0.0000**	0.050332	**0.0000**
	γ_1			0.004812	**0.0000**	0.010138	**0.0000**
	γ_2			0.680600	**0.0000**	0.693480	**0.0000**
	γ_3					0.041699	**0.0000**
Kurtosis equation	δ_0			0.448608	0.0000	0.836251	**0.0000**
	δ_1			0.000401	0.0000	0.002240	**0.0000**
	δ_2			0.810595	0.0000	0.650395	**0.0000**
	δ_3					-0.005481	**0.0000**
Log-likelihood		1687.417		2634.648		1986.751	
SIC		1655.627		2583.785		1929.531	
Ljung-Box Q-stat.							
Residuals (lag 20)		15.384	(0.754)	15.384	(0.754)	16.209	(0.704)
Residuals squared		2.2548	(1.000)	2.2042	(1.000)	2.6569	(1.000)

The basic TARCH-M with normal distribution is estimated using Quasi Maximum Likelihood (Bollerslev and Wooldridge, 1992) while other models are estimated using ML estimation. All models are estimated using Marquardt algorithm. Significant p-values are indicated by bold.

Source: Own calculations.

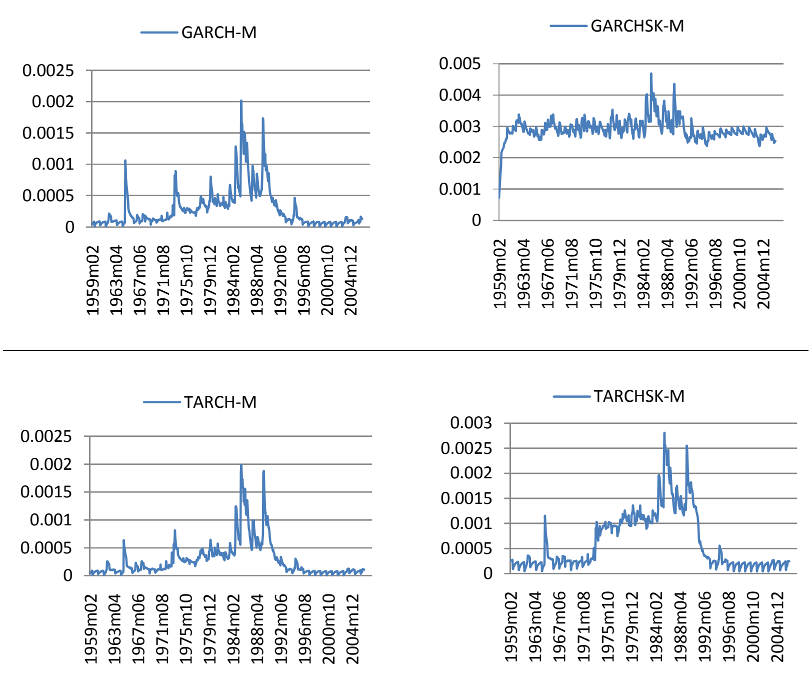
3.3. Diagnostic Tests

To compare models, I start with comparing the behaviour of the standardised residuals obtained from different models. The standardised residuals of GARCHSK-M model have a skewness of 0.73 and kurtosis of 9.19. In comparison, the standardised residuals obtained from the basic GARCH-M model have a skewness of 1.99 and excess kurtosis of 23.42. Similarly, the standardised residuals from TARCHSK-M and TARCHTSK-M have skewness and kurtosis less than the residuals from TARCH-M model. Moreover, allowing the error term to follow a GED distribution leads to the highest dispersion among all models. Thus, the standardised residuals series from models with time-varying higher order conditional moments have a lower dispersion than those resulting from time-invariant conditional skewness and kurtosis.

Another way to compare models is to evaluate the behaviour of conditional volatilities resulting from the six different models. Figure 2 shows that conditional variances obtained from models with nonconstant conditional skewness and kurtosis are smoother than conditional volatility generated by standard GARCH-M and TARCH-M models. The same conclusion could be obtained from examining the descriptive statistics of these conditional variances displayed in Table 5. It is obvious that TARCHSK-M model have the lowest Jarque-Bera statistic that measures the difference of skewness and kurtosis of the series with those from the normal distribution. Additionally, the variances of TARCHSK-M and TARCHTSK-M models have skewness and kurtosis that are lower than the variance of TARCH-M model. On the other hand, the resulting variance from GARCH-M has skewness higher than that resulting from GARCHSK-M model while the latter has higher kurtosis than the former.

The final diagnostic test is to conduct a likelihood ratio test. As it was mentioned earlier, the normal density function is nested in the G-C series expansion when s_t= 0 and k_t= 3. Therefore, the constraints imposed by the normal density function with respect to the more general density based on a G-C series expansion can be tested by applying a likelihood ratio test. To compare GARCH-M and GARCHSK-M, the value of the LR statistic is quite large resulting in rejection of the null hypothesis that the restricted density (*i.e.*, the normal density function) is the correct density. Similarly, the value of LR statistic is very high in case of comparing TARCH-M with its extensions, TARCHSK-M and TARCHTSK-M. Thus, the density that restricts the skewness and kurtosis to be time-invariant is inferior to the density that permits them to vary over time. A final remark, it was not possible to run a LR test to choose between TARCHSK-M and TARCHTSK-M as the log likelihood of the latter is less than the log likelihood of the former. Thus, the specification that allow conditional skewness and kurtosis to follow a GARCH process while the variance is derived by a TARCH process outperforms the specification that assume a TARCH structure for conditional variance, skewness and kurtosis.

Figure 2
Estimated Conditional Variances from GARCH-M, GARCHSK-M, TARCH-M and TARCHSK-M Models



Source: Author's calculation from models.

Table 4
Descriptive statistics for standardised residuals

Statistic	GARCH-M	GARCH-M (GED)	GARCHSK-M	TARCH-M	TARCHSK-M	TARCHTSK-M
Mean	0.079326	0.101916	-0.697897	0.086074	-0.133299	0.025444
Median	0.042160	0.053838	-0.732487	0.042998	-0.140971	-0.012598
Maximum	10.29504	14.64040	0.842837	10.12453	6.225940	7.245058
Minimum	-3.914970	-4.119234	-2.117273	-3.722729	-2.728351	-2.927933
Std. Dev.	0.997464	1.120017	0.296164	0.996351	0.685996	0.728253
Skewness	1.986015	3.847826	0.725296	1.857173	1.680308	1.900507
Kurtosis	23.42197	53.01585	9.193916	22.04523	19.06223	21.90398
Jarque-Bera	10406.05	61566.10	972.9378	9052.096	6474.159	8938.892
Probability	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

Source: Own calculations.

Table 5

Descriptive statistics for conditional variances

Statistic	GARCH-M	GARCH-M (GED)	GARCHSK-M	TARCH-M	TARCHSK-M	TARCHTSK-M
Mean	0.000285	0.000291	0.002915	0.000284	0.000614	0.000470
Median	0.000155	0.000127	0.002867	0.000144	0.000275	0.000260
Maximum	0.002020	0.001581	0.004689	0.001989	0.002805	0.002431
Minimum	5.14×10^{-6}	1.50×10^{-5}	0.000713	5.65×10^{-6}	3.92×10^{-6}	9.43×10^{-6}
Std. Dev.	0.000316	0.000321	0.000344	0.000324	0.000570	0.000421
Skewness	2.207727	1.546118	0.039775	2.435293	1.195691	1.700310
Kurtosis	8.542156	4.986887	10.64108	9.886839	3.804254	5.890463
Jarque-Bera	1207.174	324.7943	1403.852	1710.592	153.0380	478.8859
Probability	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

Source: Own calculations.

Table 6

Likelihood ratio tests

GARCH-M vs. GARCHSK-M		**GARCH-M vs. TARCH-M**	
Logl(GARCH-M)	1684.490	Logl(GARCH-M)	1684.490
Logl(GARCHSK-M)	2542.654	Logl(TARCH-M)	1687.417
LR	1625.386	LR	5.853
p-value	0.000	p-value	0.015
TARCH-M vs. TARCHSK-M		**TARCH-M vs. TARCHSTK-M**	
Logl(TARCH-M)	1687.417	Logl(TARCH-M)	1687.417
Logl(TARCHSK-M)	2634.648	Logl(TARCHTSK-M)	1986.751
LR	1804.257	LR	508.463
p-value	0.000	p-value	0.000

Source: Own calculations.

3.4. Forecasting Performance

The predictive power of the different models is evaluated by applying several measures reported in Table 7. The first two forecast error statistics depend on the scale of the dependent variable. Thus, they are relative measures to compare forecasts across different models. According to these criteria, the smaller the error, the better is the forecasting ability of the related model. With respect to Theil inequality coefficient, it must lie between zero and one, where zero is a sign of perfect fit. Additionally, the bias and variance proportion are indicators of how far the mean and variation of the forecast are from the mean and the variance of the actual series while the covariance proportion measures the remaining unsystematic forecasting errors. These different proportions must sum up to one where smaller bias and variation proportion refers to a better forecasts. Thus, most of the bias should be concentrated on the covariance proportion.

Table 7
Different Criterions of Forecasts Power

1. Mean square error	$MSE = \frac{1}{N}\sum_{t=T+1}^{T+N}(\hat{\pi}_t - \pi_t)^2$
2. Mean absolute error	$MAE = \frac{1}{N}\sum_{t=T+1}^{T+N}\lvert\hat{\pi}_t - \pi_t\rvert$
3. Theil inequality coefficient	$TIC = \frac{\sqrt{\frac{1}{N}\sum_{t=T+1}^{T+N}(\hat{\pi}_t - \pi_t)^2}}{\sqrt{\frac{1}{N}\sum_{t=T+1}^{T+N}\hat{\pi}_t^2} + \sqrt{\frac{1}{N}\sum_{t=T+1}^{T+N}\pi_t^2}}$
Bias Proportion	$BP = \frac{\left(\bar{\hat{\pi}} - \pi\right)^2}{\frac{1}{N}\sum_{t=T+1}^{T+N}(\hat{\pi}_t - \pi_t)^2}$
Variance proportion	$VP = \frac{(\sigma_{\hat{\pi}} - \sigma_{\pi})^2}{\frac{1}{N}\sum_{t=T+1}^{T+N}(\hat{\pi}_t - \pi_t)^2}$
Covariance proportion	$CP = \frac{2(1-r)\sigma_{\hat{\pi}}\,\sigma_{\pi}}{\frac{1}{N}\sum_{t=T+1}^{T+N}(\hat{\pi}_t - \pi_t)^2}$

Source: Pindyck and Rubinfeld (1998).

Where $\sigma_{\hat{\pi}}$, σ_{π} are the biased standard deviations of $\hat{\pi}$ and π, and *r* is the correlation between of $\hat{\pi}$ and π.

Table 8 displays the results of different measures for the out-of-sample (2007:2 to 2009:2) period. The selected sample includes two years, as inflation in actual policy conduct is likely to be forecasted in a two-year horizon. Comparing GARCH-M with GARCHSK-M model reveals that for the first criterion, GARCH-M does a better job in forecasting inflation during the out-of-sample period.

For GARCH-M, TIC of 0.66 indicates a relatively poor fit between forecasted and actual values. The bias proportion accounted for 0.17 of TIC referring to the difference between the predicted and the actual mean. Additionally, the variance proportion equals 0.615318 that is the highest amongst different models indicating the failure in tracking the actual variance path. Additionally, the bias proportion for GARCHSK-M is the highest indicating significant difference between the forecasted and actual mean. Furthermore, MAE and MSE are also the highest value, which is another indication of poor predictability of inflation. Thus, comparing GARCH-M and GARCHSK-M reveals that accounting for higher order moments did not improve inflation forecasting. In addition, allowing the error term to follow a GED distribution does not improve the forecasting also as TIC is the highest coefficient relative to other models. Moreover, both BP and VP are very high which indicates the low ability to trace both actual mean and variance. Therefore, however GARCHSK-M is the best model according to the diagnostic tests, all models based on GARCH-M specification show poor forecasting performance.

Concerning TARCH-M model, comparing it with the previous models does not imply superiority to others in most of those criteria. For TARCHSK-M where skewness and kurtosis are derived by a GARCH process, TIC of 0.52 is the lowest value compared with other models. Thus, it indicates a relatively moderate fit between forecasted and actual values. The bias proportion accounted for 0.04 of TIC referring to the similarity between the predicted and the actual mean. However, the high value of VP is an indication of the significant difference between the actual variance and the predicted from the model. Overall, this model could be regarded as the best model compared to other models as CV equals 0.44. Thus, a high proportion of the dispersion is attributed to the unsystematic forecasting errors.

According to these criteria, the model that allows nonconstant conditional skewness and kurtosis to follow a GARCH process whereas the conditional variance is derived by a TARCH process outperforms all other models. Thus, a specification that allows conditional third and fourth moments outperforms all other specification that keeps them constants.

Table 8

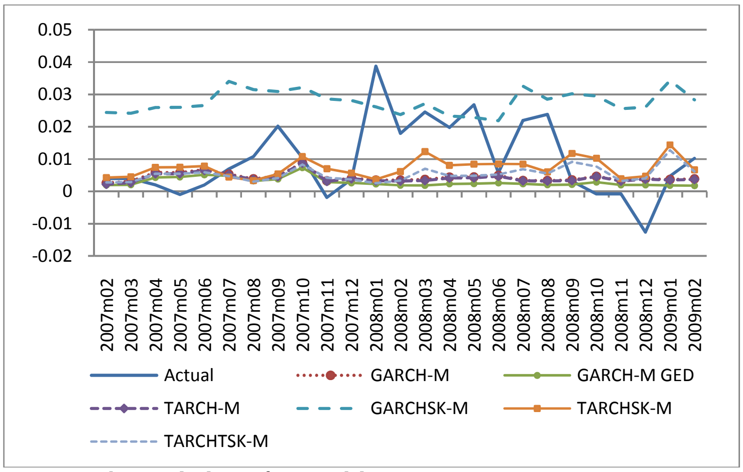
Out-of sample Forecasts power of different models[3]

	GARCH-M	**GARCH-M (GED)**	**GARCHSK-M**	**TARCH-M**	**TARCHSK-M**	**TARCHTSK-M**
MSE	0.000166 [3]	0.000184 [5]	0.000467 [6]	0.000169 [4]	0.000145 [1]	0.000161 [2]
MAE	0.009231[1]	0.009809 [5]	0.030476 [6]	0.009307 [2]	0.009391 [3]	0.009486 [4]
TIC	0.658177[4]	0.744276 [6]	0.591939 [2]	0.670647 [5]	0.527362 [1]	0.608068 [3]
BP	0.177879 [3]	0.255730 [5]	0.863351 [6]	0.189682 [4]	0.043756 [1]	0.122149 [2]
VP	0.615318 [6]	0.553804 [4]	0.067531 [1]	0.602909 [5]	0.508448 [2]	0.518346 [3]
CP	0.206802 [4]	0.190466 [5]	0.069118 [6]	0.207410 [3]	0.447795 [1]	0.359506 [2]

Numbers in brackets indicate rankings of the models where [1] indicates the best models according to the corresponding measure.
Source: Author's own calculations.

[3] It was not possible to compare the forecasting power of the models with the factor models of Stock and Watson (2002) as it requires a large set of data that is unavailable. Moreover, the conventional models of inflation forecasting, Phillips curve, requires unemployment or output data which is available only on annual basis (the quarterly series starts at the first quarter of 2003 and there is no monthly data). Therefore, to compare with the above mentioned model, the model could be applied for another advanced economy and compared with Stock and Watson (2002) models as a potential further research.

Figure 3
Actual versus predicted inflation from different models



Source: Author's calculation from models.

4. Applicability of the Model to other Economies: Evidence for Mexico

This section is devoted to check the applicability of the model to other countries conditional on that inflation data show the existence of conditional volatility. First, the appropriate mean equation has been specified according to the methodology followed in Section 4. Accordingly, inflation in Mexico is regressed on its first, fourth, and eleventh lags. Comparing the results of GARCH and TARCH models displayed in Table 9 shows that TARCH specification is superior to the GARCH model. Therefore, the analysis will extend TARCH model to allow for time-varying conditional skewness and kurtosis.

Table 9 reports the results of two basic models, GARCH and TARCH with normal distribution, and the TARCHSK model with time-varying conditional third and fourth moments. First, the shocks to inflation β_1, it is found to be significant in the standard TARCH-M model and allowing for time-varying third and fourth moments lowers its magnitude while still significant. Additionally, the persistence parameter in the variance equation is significant in all models with the highest magnitude in TARCHTSK model where both skewness and kurtosis are allowed to follow a GARCH-type process. Concerning the asymmetry term β_3 , it is found to be negative and significant in the basic TARCH model. Allowing the conditional skewness and kurtosis to follow a GARCH-type process raises β_3 and keeps it significant.

Regarding skewness equation, the persistence parameter is positive and significant although it is less than that of the variance equation. Finally, the coefficient for lagged kurtosis is the highest persistence coefficient while shocks to kurtosis have the smallest magnitude in comparison with shocks to conditional volatility and skewness.

Concerning the specification of the models, the Ljung-Box Q-statistics for the both standardised residuals and its squares are insignificant up to the twelfth lag. Thus, there is evidence that both the level and squares of standardised residuals do not exhibit any serial correlation. Finally, according to SIC criterion, TARCHSK model that allow conditional skewness and kurtosis to be time-varying is the best model.

Table 9

GARCH, TARCH, and TARCHSK Models

Mean equation: $\pi_t = \alpha_1 \pi_{t-1} + \alpha_2\pi_{t-4} + \alpha_3\pi_{t-11} + \varepsilon_t$

Variance equation: $h_t = \beta_0 + \beta_1\varepsilon_{t-1}^2 + \beta_2 h_{t-1} + \beta_3\varepsilon_{t-1}^2(\varepsilon_{t-1} < 0)$

Skewness equation: $s_t = \gamma_0 + \gamma_1\eta_{t-1}^3 + \gamma_2 s_{t-1}$

Kurtosis equation: $k_t = \delta_0 + \delta_1 \eta_{t-1}^4 + \delta_2 k_{t-1}$

Model		GARCH		TARCH		TARCHSK	
		estimate	**p-value**	**estimate**	**p-value**	**estimate**	**p-value**
	α_1	0.591692	**0.0000**	0.654241	**0.0000**	0.721093	**0.0000**
	α_2	0.116641	**0.0022**	0.135846	**0.0000**	0.118370	**0.0000**
	α_3	0.158530	**0.0000**	0.119093	**0.0000**	0.006947	**0.0000**
Variance equation	β_0	1.34×10^{-05}	**0.0000**	1.22×10^{-05}	**0.0000**	2.34×10^{-05}	**0.0000**
	β_1	0.535467	**0.0000**	0.710816	**0.0000**	0.505230	**0.0000**
	β_2	0.444531	**0.0000**	0.487616	**0.0000**	0.717362	**0.0000**
	β_3			-0.637900	**0.0000**	-0.504689	**0.0000**
Skewness Equation	γ_0					0.070435	**0.0000**
	γ_1					-0.258186	**0.0000**
	γ_2					0.599714	**0.0000**
Kurtosis Equation	δ_0					0.496505	**0.0000**
	δ_1					0.007377	**0.0000**
	δ_2					0.787612	**0.0000**
Log-likelihood		2018.093		2032.159		2162.807	
SIC		1981.830		1995.895		2126.585	
Ljung-Box Q-stat.							
Residuals (lag 12)		25.363	(0.013)	17.458	(0.133)	17.313 (0.099)	
squared (lag 12)		15.138	(0.234)	14.215	(0.287)	15.670 (0.207)	

The basic GARCH and TARCH models are estimated using Quasi Maximum Likelihood (Bollerslev- Wooldridge (1992)) TARCHSK-M model is estimated using ML estimation. All models are estimated using Marquardt algorithm. Significant p-values are indicated by bold.

Source: Own calculations.

As before, to evaluate different models, the behaviour of the standardised residuals is compared. As displayed in Table 10, the standardised residuals of TARCHSK model have the smallest standard deviation, skewness and kurtosis. Thus, the standardised residuals series from models with time-varying higher order conditional moments have a lower dispersion than those resulting from time-invariant conditional skewness and kurtosis. In addition, the behaviour of conditional volatilities resulting from different models is compared. Based on Table 11, the results indicate that conditional variance resulting from TARCHK model has the lowest skewness and kurtosis.

The final diagnostic test is to conduct a likelihood ratio test. To compare GARCH and TARCH specifications, the value of the LR statistic is quite large resulting in rejection of the null hypothesis that the restricted model (*i.e.*, GARCH) is the correct model. Similarly, the value of LR statistic is very high in case of comparing TARCH with its extension. Thus, the density that permits the skewness and kurtosis to be time-varying outperforms the density that keeps them constant.

Table 10
Descriptive statistics for standardised residuals

Statistic	**GARCH**	**TARCH**	**TARCHSK**
Mean	0.242378	0.189668	0.181784
Median	0.094101	0.059529	0.092059
Maximum	5.591658	5.638982	3.377584
Minimum	-2.422844	-2.401694	-1.855432
Std. Dev.	0.972287	0.983745	0.614443
Skewness	1.911525	1.763339	1.571154
Kurtosis	10.104180	9.133696	8.416244
Jarque-Bera	1591.871000	1224.376000	959.006200
Probability	0.000000	0.000000	0.000000

Source: Own calculations.

Table 11
Descriptive statistics for conditional variances

Statistic	**GARCH**	**TARCH**	**TARCHSK**
Mean	0.000119	0.000122	0.000224
Median	4.26×10-05	4.02×10-05	0.000122
Maximum	0.002811	0.003387	0.003065
Minimum	2.42×10-05	2.40×10-05	8.42×10-05
Std. Dev.	0.000248	0.000284	0.000297
Skewness	6.466846	7.044763	5.261333
Kurtosis	56.76229	66.05730	39.21945
Jarque-Bera	74785.37	102107.1	34793.82
Probability	0.000000	0.000000	0.000000

Source: Own calculations.

Table 12
Likelihood ratio test

GARCH vs. TARCH	
Logl(GARCH-M)	2018.093
Logl(TARCH-M)	2032.159
LR	28.131
p-value	0.000
TARCH vs. TARCHSK	
Logl(TARCH-M)	2032.159
Logl(TARCHTSK-M)	2162.807
LR	261.296
p-value	0.000

Source: Own calculations.

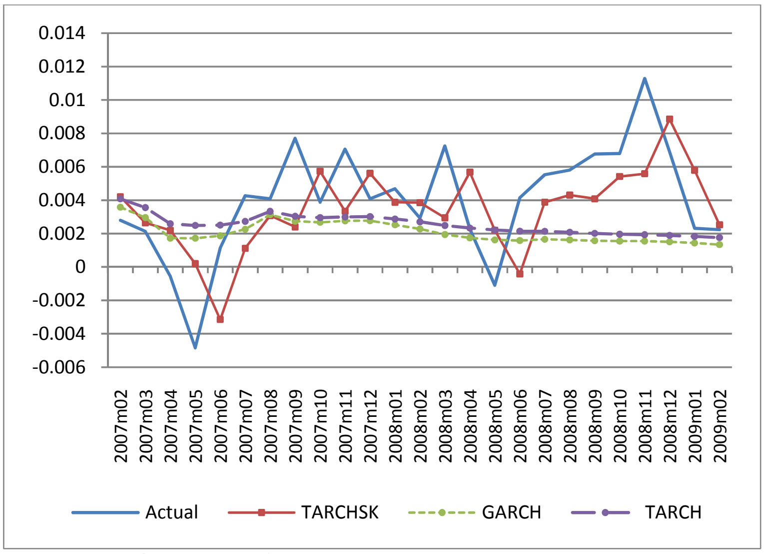
The predictive power of different models is calculated for the out-of-sample (2007:2 to 2009:2) period. The results, which are displayed in Table 13, reveal the superiority of TARCHSK model according to all the measures. Finally, Figure 4 shows the ability of TARCHSK model to trace the inflation path in comparison with standard GARCH and TARCH models. Thus, a specification that allows conditional third and fourth moments outperforms other specifications that keep them constants.

Table 13
Out-of sample Forecasts power of different models

	GARCH	TARCH	TARCHSK
MSE	1.45×10^{-05} [3]	1.35×10^{-05} [2]	9.55×10^{-06} [1]
MAE	0.003010 [3]	0.002871 [2]	0.002663 [1]
TIC	0.519435 [3]	0.474202 [2]	0.331450 [1]
BP	0.248738 [3]	0.152401 [2]	0.030950 [1]
VP	0.482294 [3]	0.531685 [2]	0.081665 [1]
CP	0.268968 [3]	0.315914 [2]	0.887385 [1]

Numbers in brackets indicate rankings of the models where [1] indicates the best models according to the corresponding measure.
Source: Own calculations.

Figure 4
Actual versus predicted inflation from GARCH, TARCH and TARCHSK Models



Source: Author´s calculation from models.

Conclusion

Given that inflation forecasts are important in the actual monetary policy conduct especially under an inflation-targeting regime, central banks must have accurate inflation forecasts. Additionally, since point forecasts provide precise predictions only if the underlying loss function is quadratic while the constraints are linear which is unrealistic in all cases, a density forecasts could help improving inflation forecasting. Therefore, the paper applied the methodology proposed by (Leon *et al.*, 2005) for modelling nonconstant conditional second, third and fourth moments to explore the full density of inflation in Egypt.

The employed methodology includes GARCH-M and TARCH-M models along with their extensions that permit conditional skewness and kurtosis to follow GARCH and TARCH structures. Additionally, a GARCH-M specification with GED distribution for the error term is modelled and compared with models that assume normality and models that assume a G-C series expansion. The results indicate that there is a significant persistence in conditional variance, skewness and kurtosis. Additionally, comparing different models by examining the behaviour of standardised residuals, conditional variances and conducting a likelihood ratio test reveal that models with nonconstant second, third and fourth moments are superior to models with time invariant volatility, skewness and kurtosis.

Applying the methodology to Mexican inflation data also support TARCHSK model that allow for time-varying conditional skewness and kurtosis. As a result, it can be concluded that monthly inflation is indeed highly asymmetric in both countries. Therefore, central banks should care about the full density of inflation in constructing their future forecasts. Finally, using annual inflation data from different counties including Egypt, Mexico, Colombia, Denmark and Finland show that annual inflation is not volatile which means the methodology applied here is not valid for them. On the other hand, checking annual inflation for both Belgium and Turkey shows evidence of the existence of ARCH effects. Thus, annual data gives mixed results regarding the possibility of applying the underlying models. The paper could be extended to check the applicability of the underlying methodology for other countries using different frequencies.

References

Balke, N. S., and M. A. Wynne (1996). "An Equilibrium Analysis of Relative Price Changes and Aggregate Inflation." *Federal Reserve Bank of Dallas Working Papers* 09.

Ball, L., and N. G. Mankiw (1995). "Relative-Price Changes as Aggregate Supply Shocks." *Quarterly Journal of Economics*, 110 (1), 161-193.

Berument, B., K. Metin-Ozcan and B. Neyapti (2001). "Modelling Inflation Uncertainty using EGARCH: An Application to Turkey." *Bilkent University Working Paper*, May.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroscedasticity." *Journal of Econometrics*, 31, 307-327.

_______ (2008). "Glossary to ARCH (GARCH)." *Centre for Research in Econometric Analysis and Time Series (CREATES)*, University of Copenhagen, Research Paper No. 49.

_______ and J. M. Wooldridge (1992). "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances." *Econometric Reviews*, 11, 143–172.

Brooks, C. (2002). "Introductory Econometrics for Finance." Cambridge University Press.

Bryan, M., and S. Cecchetti (1996). "Inflation and the Distribution of Price Changes." *NBER Working Paper 5793*.

Central Bank of Egypt (CBE) (2005). "Monetary Policy Statement." June.

Chaudhuri, K., M. Kim and Y. Shin (2011). "Forecasting Time-varying Densities of Inflation Rates: A Functional Autoregressive Approach." March, Available at SSRN: http://ssrn.com/abstract=1555257.

Engle, R. (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation." *Econometrica*, 50 (4), 267-287.

_______ (1995). "ARCH: Selected Readings." Oxford University Press.

Friedman, M. (1977). "Nobel Lecture: Inflation and Unemployment." *The Journal of Political Economy*, 85(3), 451-472.

Glosten, L. R., R. Jagannathan and D. E. Runkle (1993). "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *Journal of Finance*, 48(5), 1779-1801.

Grier, K., and M. Perry (1996). "Inflation, Inflation Uncertainty, and Relative Price Dispersion." *Journal of Monetary Economics*, 38, 391-405.

Harvey, C., and A. Siddique (1999). "Autoregressive Conditional Skewness." *Journal of Financial and Quantitative Analysis*, 34 (4), 465- 487.

Leon, A., G. Rubio and G. Serna (2005). "Autoregressive Conditional Volatility, Skewness and Kurtosis." *The Quarterly Review of Economics and Finance*, 45, 599-618.

Nelson, D. B. (1991). "Conditional Heteroscedasticity in Asset Returns: A New Approach." *Econometrica*, 59 (2), 347-370.

Noureldin, D. (2005). "Alternative Approaches to Forecasting Inflation in the Case of Egypt." Economic Research Forum 12th Annual Conference, Cairo, Dec, 19-21.

Pindyck, R. S. and D. L. Rubinfeld (1998). "Econometric Models and Economic Forecasts." 4th edition, New York: McGraw-Hill.

Rachev, S.T., S. Mittnik, F. J. Fabozzi and S. M. Focardi (2007). "Financial Econometrics: from Basics to Advanced Modelling Techniques." John Wiley, New Jersey.

Roger, S. (2000). "Relative Prices, Inflation and Core Inflation." IMF Working Papers No. 00/58.

Stock J. H. and M. W. Watson (2002). "Macroeconomic Forecasting using Diffusion Indexes." *Journal of Business and Economic Statistics*, 20, 147–162.

Tay, A.S., and K. F. Wallis (2000). "Density Forecasting: A Survey." *Journal of Forecasting*, 19, 235–254.

Wallis, K. F. (2004). "An Assessment of Bank of England and National Institute Inflation Forecast Uncertainties." *National Institute of Economic Review*, 189 (1), 64-71.

Wilson, B. (2006). "The Links between Inflation, Inflation Uncertainty and Output Growth: New Time Series Evidence from Japan." *Journal of Macroeconomics*, 28, 609-620.