UNIVERSITÀ DEGLI STUDI DI GENOVA

Dottorato di Ricerca in Economia XXX ciclo

Settore Scientifico Disciplinare: SECS-P/05 ECONOMETRIA

Essays on Financial Econometrics

Fabio Profumo

Supervisor:

Professoressa Anna Bottasso

Professore Lorenzo Trapani

Facoltà di Economia

Scuola di Scienze Sociali

Genova

Contents

List of Figures

List of Tables

Introduction

This thesis reflects my PhD studies at the University of Genova and presents the development of my research interests in the field of financial econometrics. It serves to illustrate the sheer breadth of topics associated with financial econometrics by focusing on three distinct and diverse research themes - whilst also showing the interrelationships that can exist between seemingly disparate themes. What follows are three chapters exploring price transmission dynamics, mitigation of risk, and predictive accuracy in the presence of heteroscedasticity; each chapter shows methodologies that can be applied to a wide range of financial strands including commodities market, foreign exchange and risk management. Further, empirical examples in each chapter evidence their application within each research theme.

The first chapter analyses hedge ratios and the importance of using a hedging strategy that allows financial institutions to mitigate risk and comply with enforced regulations without holding excessive amounts of funds to satisfy minimum capital requirements. The choice of an adequate hedging strategy, intended as detaining the appropriate units of spot and futures instruments to mitigate the risk of fluctuations in the price of the underlying assets, is pivotal. In my empirical exercise I consider simple methods, such as naive hedging as well as time-varying hedging obtained from a conditional OLS model, and more complicated multivariate GARCH specifications such as GJR and Exponential MGARCH with error correction term. I compare the latter methods using different measures of risk: Variance reduction with respect to the unhedged portfolio, Value at Risk and Expected Shortfall. My results show that the most simple hedging methods lead to the best performances of risk indicators.

Moving on to a wholly different theme in financial econometrics, the second chapter discusses the transmission of changes in wholesale oil prices to retail petrol prices in the UK and Italy. Asymmetric pass-though appears in various fields and most of the literature is focused on agricultural products. On fuel prices, most of the literature is concentrated on the US and the UK markets. Galeotti et al. (2003) and Grasso & Manera (2007) study the market in Italy, France, Spain, Germany and the UK but their models are estimated using a standard two-step procedure. Another way to proceed and discern between short run and long run asymmetry is to use an autoregressive distributed lag approach as in Shin, Yu & Greenwood-Nimmo (2014), which allows the modelling of dynamic asymmetries and estimate coefficients in one step by OLS differently from previous methods. My empirical study shows that short run asymmetry is detected in both the UK and the Italian markets but long run asymmetry in price transmission is only detectable in Italy.

The final chapter shows the interrelationships between seemingly dissimilar themes by drawing the heteroscedasticity theme from the first chapter and investigates tests for equal forecast accuracy, in particular, the Diebold and Mariano test for model-free forecasts. The original Diebold & Mariano (1995) paper proposes a test for equal forecast accuracy with an empirical application to exchange rates but it neglects the effect of heteroscedasticity on the estimate of the long run variance used in the test statistic. When dealing with financial time series, heteroscedasticity is an important feature to consider as ignoring it can lead to incorrect conclusions. In the case of tests for equal forecast accuracy, the phenomenon is of crucial importance, especially when the sample available for testing is small. Kiefer & Vogelsang (2005) suggest a new approach which is heteroscedasticity and serial correlation robust and proved to improve size performances of standard tests; the only drawback is that the asymptotic distribution of the test statistic is non standard and has to be obtained using simulations. To use a standard distribution, Hualde & Iacone (2015) suggest using an estimator for the long run variance in the frequency domain, paired with fixed asymptotics; the distribution of the Diebold and Mariano test under their assumptions follows a Student-t distribution. Keeping these issues in mind, I revisit the empirical application on exchange rates in the original Diebold & Mariano (1995) paper allowing for heteroscedasticity and using fixed smoothing asymptotics. My results indicate that a random walk model better forecasts the Dollar/Euro exchange rate with respect to three months forward rates.

Chapter 1

Optimal Hedge Ratios

Abstract

The aim of this chapter is to assess the most effective hedging method for value and growth S&P 500 indexes using variance reduction, Value-at-Risk and Expected Shortfall associated with every hedging strategy used. I consider a naive hedging, a time varying hedging form a conditional OLS regression and three different time varying hedging strategies from a bivariate GARCH model with a GARCH, GJR-GARCH and EGARCH specification for the conditional variance. Results show that the best performing methods in terms of risk, are obtained using a naive hedge ratio or one obtained from a conditional OLS regression. Findings are robust to the measure of risk from the time to time considered and robust before and after the subprime financial crisis.

1.1 Introduction

Hedging is an investment made to mitigate the risk of fluctuations in price of the underlying assets at maturity, it involves taking an appropriate position to the spot market so that the portfolio consisting of both spot and future contract is formed. The pivotal decision is to find the optimal hedge ratio which is defined as the ratio of the number of units traded in the futures market to the number of units traded in the spot market. According to Hatemi-J & Roca (2006) the optimal hedge ratio is defined as the quantities of the spot instrument and the hedging instrument that ensure the total value of the hedged portfolio does not change; the hedge ratio can be demonstrated as the slope coefficient in a regression of the price of the spot instrument on the price of the future (hedging) instrument. The minimum variance hedge strategies introduced by Ederington (1979) are still the most popular and widely used approaches, although this has been criticised for not taking into account the expected return which is inconsistent with the mean-variance framework. An optimal strategy increases the efficiency of risk management and minimizes the costs of hedging. In this concern, many optimal hedge ratio estimations have been conducted extensively since the introduction of future contracts. There are several techniques suggested by many researches and studies in a majority of markets: Copeland & Zhu (2006) for the foreign exchange rate and Cecchetti, Cumby & Figlewski (1988) for fixed income securities. Ghosh & Clayton (1996) and Kenourgios, Samitas & Drosos (2008) discuss hedging in the stock index and both sustain the superiority of Error Correction models in terms of risk reduction. The simplest approach to hedging is naive hedging which is a simple way to hedge risk and it consists into taking one unit of a short position of a futures contract for each unit of a long position held in the underlying asset; there is no modelling in this strategy and so no estimation and optimization. It is implemented setting the hedge ratio equal to one. This strategy is indicated when the variance of the hedged portfolio is little. To take into account the fact that spot and future price movements may be different, Ordinary Least Square regression can be used to estimate the hedge ratio (Ederington, 1979). The slope coefficient acts as hedge ratio and it is constant. Its constant nature does not allow it to capture variations in the volatility of returns, time varying distributions, serial correlation and cointegration (Poterba, Rotemberg & Summers, 1986; Bollerslev, 1986; Baillie & DeGennaro, 1990). For these reasons, time varying hedge ratios need to be considered. Conditional OLS was developed to take into account the time varying structure of the ratio; Miffre (2004) argues that it recognizes the less than perfect correlation between spot and futures prices, it captures the stochastic movements in hedge ratios arising from the predictability of returns and it is easy to estimate as it produces instant estimates of the hedge ratio. Another way of capturing stylised facts of returns is to use Autoregressive Conditional Heteroscedasticity (ARCH) models (Engle, 1982; Bollerslev, 1986) which can incorporate facts such as volatility clustering and leverage effect according to the different specifications. Engle, Lilien & Robins (1987) introduce the ARCH-M model which allows the conditional variance to affect also the mean and they used it in a context of bonds and risk premium. To model asymmetric features of volatility, Glosten, Jagannathan & Runkle (1993) develop the GJR-GARCH which include leverage effect in the conditional variance specification. To further improve the conditional variance specification, Nelson (1991) suggests Exponential GARCH which not only can reproduce leverage effect but also persistence of shocks on volatility. In addition, due to the fact that the conditional variance is expressed as a logarithmic function, it cannot be negative even if parameters in its specification are negative. Another well-known feature of financial time series is long run cointegration and it is fairly common between spot and future rates. Ghosh & Clayton (1996) extend the usual hedging method with an Error Correction Model while Kroner & Sultan (1993) use a bivariate GARCH model with Error Correction to estimate the hedge ratio for currencies. In addition, using a bivariate GARCH keeps into account comovements of spot and future rates in the covariance matrix making hedging more precise. Bollerslev, Engle & Wooldridge (1988) extended a univariate GARCH to a specification with vectorised conditional covariance matrix, Baillie & Myers (1991) use this VECH-GARCH for covariance matrix of commodities' spot and future prices. Also the BEKK-GARCH model by Baba, Engle, Kraft & Kroner (1990) is widely used as the positive definiteness of the conditional covariance matrix is defined by construction. To overcome estimation complexity of multivariate model, Bollerslev (1990) introduces Conditional Constant Correlation models which assume the conditional correlation of different assets is invariant; Engle (2002) relaxed this assumption suggesting Dynamic Conditional Correlation models (DCC) in which the estimation is divided in two steps: general parameters first and correlation parameters later. A bivariate ECM-GARCH is used by Park & Jei (2010) to reflect the problem of heteroscedasticity; they use a bivariate Dynamic Conditional Correlation GARCH to model conditional covariance matrix of errors. In order to incorporate effects of previous positive and negative shocks separately in the conditional covariance matrix, the Glosten, Jagannathan & Runkle (1993)[GJR] specification is used in the individual GARCH process and the conditional correlation. More complex multivariate GARCH specifications are available like a multivariate EGARCH from Batten & In (2006). Other ways of modelling volatility are available: another parametric approach is Stochastic Volatility (SV) models proposed by Taylor (1986) and Clark (1973) in which volatility is modelled as unobservable or, in a non-parametric setting, Realised Volatility (RV) models by Andersen, Bollerslev & Diebold (2002) are

a proxy for non-observable integrated volatility which proved to be reliable with high frequency intraday data. In particular, Bkedowska-Sojka & Kliber (2010) maintain RV models work better with high frequency data otherwise, GARCH or SV models have similar performances in estimating volatility. Also GARCH and SV models have comparable performances as noted in Shephard (1996), Kim, Shephard & Chib (1998), So, Lam & Li (1999), Ruiz, Pena & Carnero (2001) and Lehar, Scheicher & Schittenkopf (2002) but SV specifications have the disadvantage of a complex and difficult estimation. Moreover, Ammann, Skovmand & Verhofen (2009) suggest SV and IV models have a tendency to overestimate volatility and risk making these models not suitable for hedging as this feature may lead to more expensive operations for financial institutions.
With such a variety of hedging techniques, there is the need to evaluate their effectiveness. For this purpose, Ederington (1979) suggests looking at the variance reduction of a hedging method with respect to the unhedged portfolio. However, this measure of risk does not take into account that variance is influenced by both positive and negative events while investors are only concerned with negative events. Other measures of risk based on the loss distribution tails may be more appropriate such as Value at Risk by Jorion (2000) and Expected Shortfall by Artzner, Delbaen, Eber & Heath (1999).

In this chapter, I use all these three hedging evaluation methods to compare the effectiveness of naive hedging, the conditional OLS time varying hedge by Miffre (2004) and time varying hedge from DCC-MGARCH with ECM with three different specifications for the conditional variance: GARCH, GJR-GARCH and Exponential GARCH on daily data of Value and Growth S&P500 indexes from 7 November 1995 to 5 December 2014. Results show that the most complicated models such as MGARCHs are beaten by naive hedging or conditional OLS hedging confirming results from Lien (2009) and Lien (2010). The remainder of this chapter is organised as follows: section two describes models used, section three illustrates hedging comparison methods, section four describes the empirical exercise and section five concludes.

1.2 Models Specification

A straightforward hedging strategy is the static hedge ratio β given by the OLS regression

$$\Delta s_t = \alpha + \beta \Delta f_t + \epsilon_t \tag{1.1}$$

where S_t and F_t are the natural logarithm of the spot and future prices at time t respectively, $\Delta s_t = \ln(S_t/S_{t-1}) = s_t - s_{t-1}$ and $\Delta f_t = \ln(F_t/F_{t-1}) = f_t - f_{t-1}$ are compounded returns for the spot rate and the future rate respectively. The slope coefficient β represents the static hedge ratio that, according to Lien (2009) and Lien (2010), has the best performances in minimising the unconditional variance.

Although it is simple, this type of hedging strategy does not take into account the fact that the variance-covariance matrix of returns is likely to change over time. In this concern, I use several types of time varying hedge ratio which aim to minimise the conditional variance instead of the unconditional variance.

The first one, it is a simple but efficient hedge ratio that has been proposed by Miffre (2004): the conditional OLS minimum variance hedge ratio. This is obtained by the regression

$$\Delta s_t = \alpha + \beta_{t-1} \Delta f_t + \epsilon_t \tag{1.2}$$

where β_{t-1} depends on the lagged basis $s_{t-1} - f_{t-1}$

$$\beta_{t-1} = \beta_0 + \beta_1(s_{t-1} - f_{t-1}) \tag{1.3}$$

Combining (1.2) and (1.3), I get a non-linear model:

$$\Delta s_t = \alpha + \beta_0 \Delta f_t + \beta_1 z_{t-1} \Delta f_t + \epsilon_t \tag{1.4}$$

where $z_{t-1} = s_{t-1} - f_{t-1}$ is the interaction variable which is supposed to pick up the variation in the hedge ratio related to changing economic conditions. The static model in (1.1) is nested in (1.2) as, when there is no new information available, the time varying hedge ratio will reduce to a static ratio.
Another way of keeping into account the time varying characteristic of the covariance matrix is to use conditional heteroscedasticity models. Baillie & Myers (1991), Kroner & Sultan (1990) and Park & Switzer (1995) among others, use MGARCH models and obtain reliable hedge ratios. Also, ignoring the cointegration relationship between spot and future returns may lead to erroneous hedge ratios as argued by Ghosh (1993a), Ghosh (1993b) and Lien (1996) among others. For these reasons, I use different specifications of autoregressive conditional heteroscedasticity models with an error correction term.

I first consider a bivariate DCC-MGARCH as in Park & Jei (2010)

$$\begin{bmatrix} \Delta f_t \\ \Delta s_t \end{bmatrix} = \begin{bmatrix} \mu_f \\ \mu_s \end{bmatrix} + \sum_{i=1}^{m} \Gamma_i \begin{bmatrix} \Delta f_{t-i} \\ \Delta s_{t-i} \end{bmatrix} + \begin{bmatrix} \mu_f \\ \mu_s \end{bmatrix} \upsilon_{t-1} + \begin{bmatrix} \epsilon_{f,t} \\ \epsilon_{s,t} \end{bmatrix} \tag{1.5}$$

where $\upsilon_{t-1} = s_{t-1} - f_{t-1}$ is the error correction term.

I assume that the covariance matrix Σ_t of $\boldsymbol{\epsilon_t} = \begin{bmatrix} \epsilon_{f,t} \\ \epsilon_{s,t} \end{bmatrix}$ follows a bivariate Dynamic Conditional Correlation GARCH, $\Sigma_{t-1} = D_{t-1} R_{t-1} D_{t-1}$ where

$$D_{t-1} = \begin{bmatrix} \sigma_{1,t-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{m,t-1} \end{bmatrix}, \quad R_{t-1} = \begin{bmatrix} 1 & \rho_{12,t-1} & \cdots & \rho_{1m,t-1} \\ \rho_{21,t-1} & 1 & \cdots & \rho_{2m,t-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m1,t-1} & \rho_{m2,t-1} & \cdots & 1 \end{bmatrix}$$

$\sigma^2_{i,t-1} = Var(r_{it}|\Omega_{t-1})$ denotes the conditional variance of the ith return and $\rho_{ij,t-1} = \rho_{ji,t-1} = (Cov(r_{it}, r_{jt}|\Omega_{t-1}))/(\sigma_{i,t-1}\sigma_{j,t-1})$ is the dynamic conditional correlation as in Engle (2002).

I first estimate the following GARCH(1,1) model for the conditional variance $\sigma^2_{i,t-1}$:

$$\sigma^2_{i,t-1} = \bar{\sigma}^2_i(1 - \lambda_{1i} - \lambda_{2i}) + \lambda_{1i}\sigma^2_{i,t-2} + \lambda_{2i}r^2_{i,t-1} \tag{1.6}$$

where $\bar{\sigma}^2$ is the unconditional variance of the ith return, and I estimate the (i,j)th conditional correlation as

$$\tilde{\rho}_{ij,t-1}(\phi) = \frac{q_{ij,t-1}}{\sqrt{q_{ii,t-1}q_{jj,t-1}}} \tag{1.7}$$

where $q_{ij,t-1} = \bar{\rho}_{ij}(1 - \phi_1 - \phi_2) + \phi_1 q_{ij,t-2} + \phi_2 \tilde{r}_{i,t-1}\tilde{r}_{j,t-1}$, $\phi_1 + \phi_2 < 1$ and $\bar{\rho}_{ij}$ is the (i,j)th unconditional correlation and $\tilde{r}_{i,t-1}$ is the standardised return.

To allow for asymmetric effects in the conditional variance of positive and negative shocks, I also use the GJR-GARCH specification (Glosten, Jagannathan & Runkle,

1993) for the conditional variance setting

$$\sigma^2_{i,t-1} = \alpha_0 + \gamma_1 \sigma^2_{t-2} + \gamma_2 \epsilon^2_{t-2} + \tau I_{t-2} \epsilon^2_{t-2} \tag{1.8}$$

where $I_t = 1$ when $\epsilon_t < 0$ and 0 otherwise. I also consider the Exponential GARCH by Nelson (1991) similarly to Batten & In (2006) setting

$$\sigma^2_{i,t-1} = \sigma^{2\gamma}_{t-2} e^{\alpha_0 + \alpha_1 g(\epsilon_{t-2})} \tag{1.9}$$

where $g(\epsilon_t) = \theta\epsilon_t + \rho[|\epsilon_t| - E(|\epsilon_t|)]$. γ measures the persistence of volatility, ρ the magnitude of the shock and θ the leverage effect.
The time varying hedge ratio is then obtained as

$$\beta_{t-1} = \frac{\hat{\sigma}_{sf,t}}{\hat{\sigma}^2_{f,t}} = \hat{\rho}_{12} \frac{\hat{\sigma}_{s,t}}{\hat{\sigma}_{f,t}} \tag{1.10}$$

where $\hat{\sigma}^2_{s,t}$ and $\hat{\sigma}^2_{f,t}$ are estimated conditional variances of the spot return and the future return respectively from the various GARCH specifications and $\hat{\rho}_{12}$ is the estimated conditional correlation.

1.3 Hedging Effectiveness

Assuming the conditional variance of the expected return of the portfolio can be evaluated as

$$h_{pt} = VAR(\Delta s_t) + \beta^2_{t-1} VAR(\Delta f_t) - 2\beta_{t-1} COV(\Delta s_t; \Delta f_t), \tag{1.11}$$

I analyse the hedging effectiveness of the naive and the time-varying hedge ratios compared to the unhedged position ($\beta = 0$). The analysis is performed for both value and growth indexes.

A common way to assess hedging performance is variance reduction relative to the unhedged portfolio (Ederington, 1979). Given $VAR(P^h)$ the variance of the hedged portfolio and $VAR(P^s)$ the variance of the unhedged portfolio, the Variance Reduction can be expressed as

$$VR = 1 - \frac{VAR(P^h)}{VAR(P^s)}. \tag{1.12}$$

However, considering the variance a measure of risk has some disadvantages as the variance takes into account also positive events, not only the bad ones and if the underlying distribution is not normal, it is not the proper scale. In this light, I also consider two measures based on the tail of the loss distribution: Value at Risk (VaR) (Jorion, 2000) and the Expected Shortfall (ES) (Artzner, Delbaen, Eber & Heath, 1999) which are commonly employed in determining the capital requirements of financial institutions.

Both tail measures are deemed coherent in the sense that for any two loss random variables X and Y, the risk measure η satisfies the following properties:

i subadditivity: $\eta(X + Y) \leqslant \eta(X) + \eta(Y)$,

ii monotonicity: if $X \leqslant Y$ almost surely, then $\eta(X) \leqslant \eta(Y)$,

iii positive homogeneity: for $c > 0$, $\eta(cX) = c\eta(X)$,

iv translation invariance: for $c > 0$, $\eta(c + X) = c + \eta(X)$.

VaR represents the amount a position could decline in a given period associated with a given probability

$$VaR_\alpha = \inf\{l \in R : P(L > l) \leqslant \alpha\}. \quad (1.13)$$

The VaR of the portfolio at the significance level α is given by the smallest number l such that the probability the loss L exceeds l is not larger than α.

Similarly to Brooks, Henry & Persand (2002), I compare types of hedging using VaR assuming the loss distribution is Normal and, in this case, the VaR over a one day period can be calculated as

$$VaR_\alpha = \Phi^{-1}(1-\alpha)\sigma_t \quad (1.14)$$

where $\Phi^{-1}(1-\alpha)$ is the inverse of the cumulative density function of L which, for $\alpha = 0.05$, in the case of a Normal distribution, is 1.645.

VaR is commonly used and it is coherent for normally distributed loss but it does not describe the tail behaviour of the loss random variable L therefore, a measure like the Expected Shortfall, that is more sensitive to the shape of the loss distribution in the tail, should be preferred.

The ES, also known as Tail Value at Risk or Conditional Value at Risk, is the expected loss of a financial position after a fatal event

$$ES_\alpha = E(L|L < VaR) = \frac{\int_{-\infty}^{VaR} lf(l)dl}{P(L < VaR)}. \quad (1.15)$$

ES is the expected loss of L given that L exceeds its VaR. In case the loss distribution is Normal, the ES over a one day period can be calculated using

$$ES_\alpha = \frac{\phi(\Phi^{-1}(1-\alpha))}{\alpha}\sigma_t \tag{1.16}$$

where ϕ is the density function of a Normal random variable.

1.4 Empirical Exercise

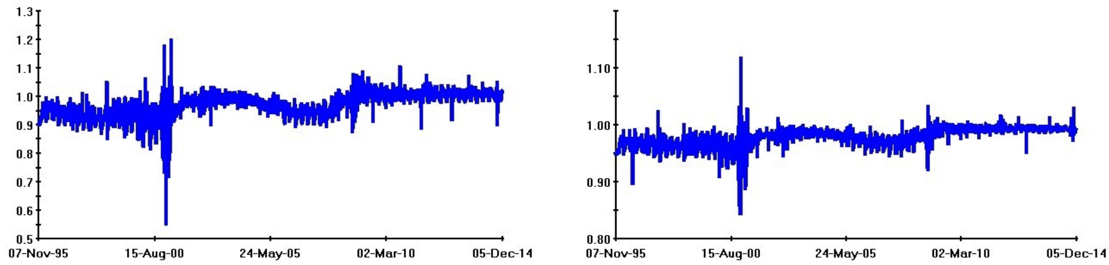
I consider daily data over the period 7 November 1995 to 5 December 2014 for the S&P500 value and growth spot index and the S&P500 value and growth future index, a total of 4981 observations. Value stocks are measured by book to market ratio, cash flow to market ratio, sales to market ratio and dividend yield, while growth stocks are measured by 5-year earnings (sales) per share growth rate and 5-year internal growth rate.

First, I use the OLS model in (1.2), for both value and growth indexes. Regression shows the intercept is not significant while, the hedge ratio is highly significant and very close to one so the null of no hedging is strongly rejected. R squared is very high as expected in this time series regressions. Tests for the null of no serial correlation (Durbin & Watson, 1950, 1951), unconditional homoscedasticity (Koenker, 1981) and normality (Jarque & Bera, 1980) are performed and their respective null hypotheses are all rejected.

I also perform a residual based Engle-Granger-Dickey-Fuller cointegration test and then I test for naive hedging ($H_0 : \beta = 1$) and for no hedging ($H_0 : \beta = 0$). As hedge ratios obtained from the regression are very close to one (Figure 1.1), I test whether a naive hedging may be used with a Wald test for $H_0 : \beta = 1$ and this hypothesis is strongly rejected, so data suggests there is no need for naive hedging.

Figure 1.1: Time-varying of hedge ratio from the conditional OLS model in (1.2)

(a) Value Index (b) Growth Index

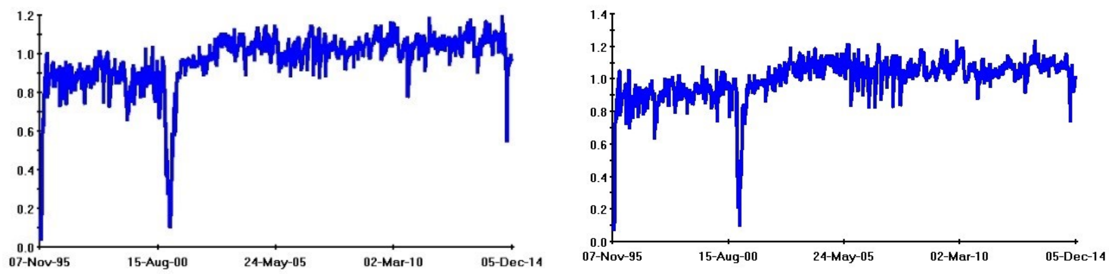


Estimating the non-linear regression in (1.4), I find the intercept is again not significant but both β_0 and β_1 are highly significant and they will both be used to construct the time-varying hedging ratio. Test statistics again reject the null of no serial correlation, homoscedasticity and normality.

Then, I generate hedging ratios from bivariate conditional heteroscedasticity models with error correction as in (1.5) with GARCH, GJR-GARCH and EGARCH specifications for the conditional variance. The hedge ratio is obtained by $\beta_{t-1} = \hat{\sigma}_{S,t}/\hat{\sigma}^2_{F,t} = \hat{\rho}_{12}(\hat{\sigma}_{S,t}/\hat{\sigma}_{F,t})$ and shown in figures 1.2, 1.3 and 1.4.

Figure 1.2: Time-varying hedge ratio from MGARCH model

(a) Value Index (b) Growth Index



All the generated hedging ratios seem to be close to one and decrease right before a crisis and then rise after the crisis as I can see during the dot com crisis in

2000 and subprime crisis in 2008. During the dot com crisis, hedge ratios became more than one so hedging more than the actual position. The ratio exhibits high volatility and quite persistence; it appears to have the same path as the basis. A possible explanation for this is that if volatility of the future is much smaller than the volatility of the spot, we might need a lot more futures to minimize the high spot variance. Hedge ratios from both Value and Growth indexes essentially show the same behaviour.

Figure 1.3: Time-varying hedge ratio from GJR - MGARCH model

(a) Value Index (b) Growth Index

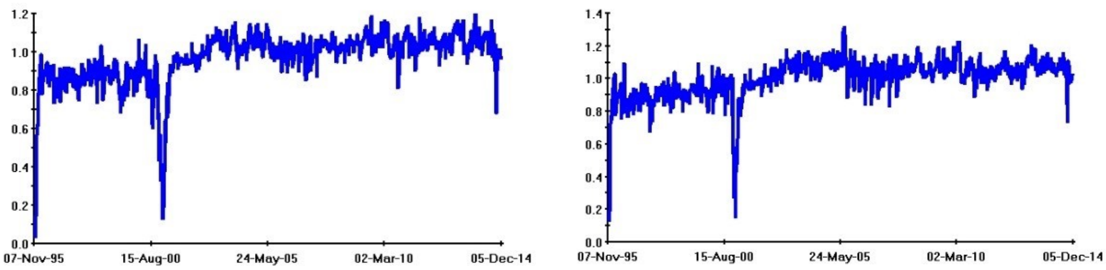
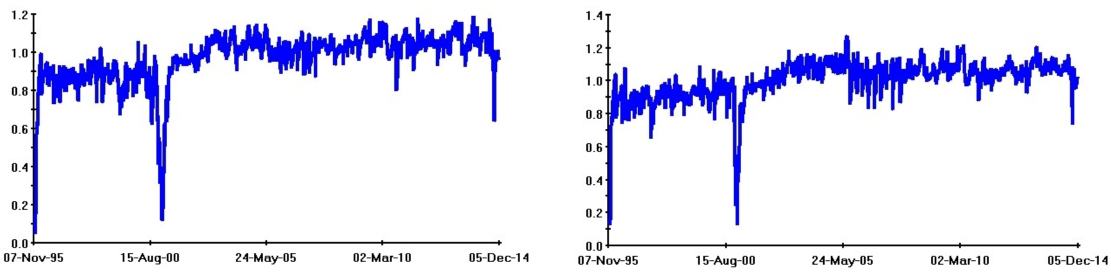


Figure 1.4: Time-varying hedge ratio from Exponential - MGARCH model

(a) Value Index (b) Growth Index



Once obtained all the hedge ratios, I can compare the effectiveness for every type of hedging considering variance reduction, VaR and ES calculated at 5% using a Normal distribution, 1 day horizon; tables 1 and 4 show these results. In addition to full sample comparison, I also compare hedging results dividing the sample in two

Table 1.1: Value Index - All sample comparison table. Figures in bold indicate the best result each row. VaR and ES are calculated over a one day period.

	No hedging $\beta = 0$	Naïve Hedging $\beta = 1$	Time-varying from conditional OLS	Time-varying from MGARCH	Time-varying from GJR MGARCH	Time-varying from E MGARCH
Return	**1.11E-02**	2.37E-03	6.68E-03	6.01E-03	6.02E-03	6.02E-03
Volatility	7.83E-03	**2.66E-03**	2.69E-03	3.13E-03	3.11E-03	3.12E-03
VR	-	**6.60E-01**	6.57E-01	6.00E-01	6.03E-01	6.01E-01
VaR	-1.29E-02	**-4.37E-03**	-4.42E-03	-5.16E-03	-5.12E-03	-5.13E-03
ES	-1.62E-02	**-5.48E-03**	-5.54E-03	-6.46E-03	-6.42E-03	-6.44E-03

sub-sections: before (07/11/1995-29/06/2007) and after the subprime crisis of mid-2007 (02/07/2007-05/12/2014). Tables 2, 3, 5 and 6 show results for sub-samples.

Table 1.2: Value Index - before crisis comparison table. Figures in bold indicate the best result each row. VaR and ES are calculated over a one day period.

	No hedging $\beta = 0$	Naïve Hedging $\beta = 1$	Time-varying from conditional OLS	Time-varying from MGARCH	Time-varying from GJR MGARCH	Time-varying from E MGARCH
Return	**1.11E-02**	2.37E-03	2.60E-03	3.41E-03	3.50E-03	3.47E-03
Volatility	7.83E-03	**2.66E-03**	2.69E-03	3.44E-03	3.26E-03	3.34E-03
VR	-	**6.60E-01**	6.56E-01	5.60E-01	5.84E-01	5.74E-01
VaR	-1.29E-02	**-4.37E-03**	-4.43E-03	-5.66E-03	-5.36E-03	-5.49E-03
ES	-1.62E-02	**-5.48E-03**	-5.55E-03	-7.10E-03	-6.72E-03	-6.88E-03

The type of hedging that generates less risk for value index is always the naive hedging, although the highest return is obtained with no hedging. The naive hedging is the one that generates the smallest return compared to other time-varying approaches. Considering this aspect, the best hedging strategy relative to risk/return trade off, could be the conditional OLS one as it provides almost three times the return of the naive case with similar level of risk. A resembling behaviour is shown in subsamples but without the return gain as in the full sample.

Table 1.3: Value Index - after crisis comparison table. 02/07/2007-05/12/2014. Figures in bold indicate the best result each row. VaR and ES are calculated over a one day period.

	No hedging $\beta = 0$	Naïve Hedging $\beta = 1$	Time-varying from conditional OLS	Time-varying from MGARCH	Time-varying from GJR MGARCH	Time-varying from E MGARCH
Return	**1.11E-02**	2.37E-03	2.20E-03	-4.17E-03	-4.66E-03	-4.46E-03
Volatility	7.83E-03	**2.66E-03**	2.68E-03	2.88E-03	2.88E-03	2.88E-03
VR	-	**6.60E-01**	6.58E-01	6.32E-01	6.32E-01	6.32E-01
VaR	-1.29E-02	**-4.37E-03**	-4.41E-03	-4.74E-03	-4.75E-03	-4.74E-03
ES	-1.62E-02	**-5.48E-03**	-5.53E-03	-5.94E-03	-5.95E-03	-5.94E-03

Table 1.4: Growth Index - All sample comparison table. Figures in bold indicate the best result each row. VaR and ES are calculated over a one day period.

	No hedging $\beta = 0$	Naïve Hedging $\beta = 1$	Time-varying from conditional OLS	Time-varying from MGARCH	Time-varying from GJR MGARCH	Time-varying from E MGARCH
Return	1.32E-02	1.89E-03	2.60E-02	**2.63E-02**	**2.63E-02**	**2.63E-02**
Volatility	1.59E-02	3.05E-03	**3.04E-03**	3.43E-03	3.41E-03	3.42E-03
VR	0.00E+00	8.08E-01	**8.09E-01**	7.84E-01	7.85E-01	7.85E-01
VaR	-2.61E-02	-5.01E-03	**-5.00E-03**	-5.65E-03	-5.62E-03	-5.63E-03
ES	-3.27E-02	-6.29E-03	**-6.26E-03**	-7.08E-03	-7.04E-03	-7.05E-03

Table 1.5: Growth Index - before crisis comparison table. 07/11/1995-29/06/2007. Figures in bold indicate the best result each row. VaR and ES are calculated over a one day period.

	No hedging $\beta = 0$	Naïve Hedging $\beta = 1$	Time-varying from conditional OLS	Time-varying from MGARCH	Time-varying from GJR MGARCH	Time-varying from E MGARCH
Return	1.32E-02	1.89E-03	**2.59E-02**	2.58E-02	2.58E-02	2.58E-02
Volatility	1.59E-02	3.05E-03	**3.04E-03**	3.51E-03	3.47E-03	3.48E-03
VR	-	8.08E-01	**8.09E-01**	7.79E-01	7.82E-01	7.81E-01
VaR	-2.61E-02	-5.01E-03	**-5.00E-03**	-5.77E-03	-5.70E-03	-5.72E-03
ES	-3.27E-02	-6.29E-03	**-6.26E-03**	-7.23E-03	-7.15E-03	-7.17E-03

Table 1.6: Growth Index - after crisis comparison table. 02/07/2007-05/12/2014. $*$ indicates the smallest number each row. VaR and ES are calculated over a one day period.

	No hedging $\beta = 0$	Naïve Hedging $\beta = 1$	Time-varying from conditional OLS	Time-varying from MGARCH	Time-varying from GJR MGARCH	Time-varying from E MGARCH
Return	1.32E-02	1.89E-03	2.61E-02	2.70E-02	**2.71E-02**	**2.71E-02**
Volatility	1.59E-02	3.05E-03	**3.04E-03**	3.32E-03	3.33E-03	3.33E-03
VR	-	8.08E-01	**8.09E-01**	7.91E-01	7.90E-01	7.90E-01
VaR	-2.61E-02	-5.01E-03	**-5.00E-03**	-5.46E-03	-5.48E-03	-5.47E-03
ES	-3.27E-02	-6.29E-03	**-6.27E-03**	-6.84E-03	-6.87E-03	-6.86E-03

Looking at the growth index, instead, the minimum risk is always obtained using a time-varying hedge ratio from a conditional OLS model. This finding is confirmed across subsamples. In terms of return, for both full sample and second sub sample, GARCH specifications lead to higher returns but the difference is negligible. In the first sample instead, the conditional OLS method also gives the best return. However, all returns are very close to zero as expected because hedging should just cover from risk and not generate revenue.

1.5 Conclusions

In this chapter, I assess which model, among a naive hedging, time-varying conditional OLS hedging or a time-varying MGARCH with GARCH, GJR-GARCH and EGARCH specifications for the conditional variance, provides the best hedging strategy for daily S&P500 Value and Growth indexes taken from 7 November 1995 to 5 December 2014 in terms of Variance Reduction, Value-at-Risk and Expected Shortfall. Results from the Value index suggests a naive hedge ratio is the one that has the best variance reduction, the smallest Value-at-risk and Expected Shortfall values. Very similar results are obtained with the time-varying hedge ratio from a conditional OLS model. In the case of the Growth index, the best performances are always obtained with the time-varying conditional OLS hedge ratio confirming findings from Lien (2009) and Lien (2010) that more complex models are not always the best hedging strategy. Results are robust across subsamples before and after the subprime financial crisis.

Chapter 2

Asymmetric pass-through of the Gasoline Prices: ‘Rockets and Feathers’

Abstract

The aim of this chapter is to assess whether positive and negative variations in wholesale prices of crude oil have different effects on gasoline retail prices (Rockets and Feathers hypothesis) in the UK and Italy. In order to do so, I use both symmetric and asymmetric OLS and ARDL regressions, I test for cointegration using the PSS Bounds test and then I perform several Wald tests for symmetry in the long run and short run and for complete pass-through. For the UK, there seems to be asymmetries only in the short run while in Italy asymmetries persist also in the long run. As a result, price changes in wholesale prices are passed to pump prices differently

according to their sign of variations.

2.1 Introduction

A common concern around consumers and regulatory bodies regards price changes and their consequences, in particular the transmission of changes in wholesale prices to retail prices. Peltzman (2000) observes that retail prices, and in general all output prices, tend to respond faster to increases in wholesale prices than decreases and this phenomenon can be observed in various fields such as financial markets, energy, agricultural and alimentary sectors. When retail prices react more fully or faster to an increase in wholesale prices, the asymmetry is called positive, conversely, when retail prices react faster to a decrease in wholesale prices the asymmetry is negative.

Most of the literature about price transmission is focused on agricultural economics starting from Heien (1980) which analyses the role of mark up costs in the food industry, Ward (1982) that studies fresh vegetables markets and Freebairn (1984) that investigate farm products to more recent studies such as Zachariasse & Bunte (2003) about meats and potatoes or Girapunthong, VanSickle & Renwick (2003) about the tomatoes market in the US among others. A survey in this field is provided by Meyer & von Cramon-Taubadel (2004) while a wider and thorough analysis in terms of markets involved and study methods can be found in Frey & Manera (2007). The first attempt to empirically investigate asymmetric pass-through tracks back to Farrell (1952) which considers a logarithmic transformation of the demand function for tobacco that ends up being the classic asymmetric specification still used nowadays. Since then, several empirical studies, mostly about agriculture, started populating literature bringing with them theoretical models for asymmetries. A strand of literature focuses on Autoregressive Distributed Lag (ARDL) models. In an ARDL model, the dependent variable depends on its own lags, the autoregressive part, and on a series of variables, contemporaneous and lagged, which represents the distributed lag part. If series involved are stationary, parameters can be consistently estimated with OLS, while, in case of non-stationary time series, models needs to be estimated using first differences. A well-known approach for modelling non-stationary series is the one suggested by Engle & Granger (1987) which develop an Error Correction Model (ECM) based on the idea that the explanatory variables have a non-linear impact on the dependent variable. The Engle-Granger approach consists in two steps: first, the equilibrium relation is estimated and tested for cointegration, second, estimate the ECM regression in which all variables are expressed in first differences except for the stationary residuals obtained from the first step. However, this method assumes only a symmetric long run relationship not allowing for any asymmetry in the long run and a single step estimation may be more efficient and improve performances as it does not suffer from any estimation uncertainty or errors arising from estimation of the long run cointegrating relation in a separate step. More recently, Shin, Yu & Greenwood-Nimmo (2014) advance the asymmetric Autoregressive Distributed Lag approach by combining the short run adjustment asymmetry with the long run reaction asymmetry. Furthermore, they propose to employ asymmetric dynamic multipliers, thereby providing a flexible means of modelling dynamic asymmetries and investigating the traverse between the short and the long run. Their empirical findings on the asymmetric retail gasoline price adjustments in Korea, clearly demonstrate that the imposition of invalid long run and or short run restrictions is likely to yield profoundly misleading results. One of the principle benefits of this approach is that, quite unlike Markov-switching or smooth transition models, it is easily estimable by standard OLS in one step.

Along with theoretical models, the empirical literature have always tried to find an economic reason for asymmetries and several causes have been identified: Eckert (2013) suggests market power and oligopolistic pricing practices which may lead to collusion and so, to asymmetric price response. This reason is supported by other authors such as Borenstein & Shepard (1996), Borenstein, Cameron & Gilbert (1997), Verlinda (2008), Balmaceda & Soruco (2008) and Deltas (2008) among others. Eckert (2013) also finds consumer search can be accountable for asymmetries; in this concern Johnson (2002), Yang & Ye (2008), Tappata (2009) and Lewis (2011) provide evidence that the ease and readiness of consumers to search influences the asymmetric response of prices. Meyer & von Cramon-Taubadel (2004) also suggest adjustment and menu costs can play a role in price transmission with several studies from Heien (1980), Ward (1982), Bailey & Brorsen (1989) and Peltzman (2000) among others. Bailey & Brorsen (1989) also add that asymmetric information can lead to asymmetric price transmission referring to the US broiler market. Other reasons include the shape of the demand function (Bonnet & Villas-Boas, 2016), psychological pricing points (Blinder, Canetti, Lebow & Rudd, 1998), production and inventory capacity and costs (Borenstein & Shepard, 2002), the non-equivalence of demand vs. supply side shocks (Kinnucan & Forker, 1987; Cramon-Taubadel, 1998), tax legislation (Greenwood-Nimmo & Shin, 2013) and, Eckert (2002), Noel (2009) and Chesnes (2016) argue that, when a market exhibits an Edgeworth cycle, the effect of a shock is reflected in retail prices in a different way according to the position of the cycle. Among these reasons, only some are more specific to the oil sector such as market power, customer search behaviour, price cycles and tax regulations while the remaining are common in the agriculture sector.

In the energy context, positive asymmetric pass-through of prices is known as the rockets and feathers hypothesis. Prices may exhibit asymmetry both in the short run, if the different reaction is only temporary, and in the long run if the initial type of adjustment persists. Many authors started focusing on oil price asymmetries since the first report from the US Monopolies and Mergers Commission in 1965. Bacon (1991), first, develops a quadratic quantity adjustment model to test whether the speed of adjustment in UK retail prices to cost changes in crude oil is rapider when costs raise than when they fall. He finds that increases in the product price are transmitted faster than reductions. Around the same time, Karrenbrock (1991) assess the market in the 80s obtaining mixed results. Manning (1991) uses an Error Correction Model (ECM) that allows for asymmetry and he finds that weak and non-persistent asymmetry is absorbed in few months.

Borenstein, Cameron & Gilbert (1997) analyse the US market using weekly data and they maintain that retail gasoline prices react more quickly to increases in the crude oil than to decreases.

Galeotti, Lanza & Manera (2003) carry out a comparison of five European countries and they employ asymmetric ECM to distinguish between short and long run asymmetries. They find several differences in adjustment speed and short run responses to price changes.

Bachmeier & Griffin (2003) consider daily data to obtain more reliable estimates if gasoline prices respond almost immediately to crude price changes and, in contrast with Borenstein et al. (1997), they find no evidence of asymmetry in wholesale gasoline prices.

Grasso & Manera (2007) use monthly data for several European countries and they find mixed results according to the model used. In their case, only a Threshold Autoregressive Error Correction Model suggests asymmetry. Murry & Zhu (2008)

investigates the US natural gas market and their dynamic model shows evidence of short term asymmetric effects.
Other studies about gasoline prices include Duffy-Deno (1996), Balke, Brown & Yucel (1998), Deltas (2008), Kaufmann & Laskowski (2005) and Radchenko (2005) among others.

In this chapter, I consider monthly data from January 1999 to March 2008 for gasoline wholesale and retail prices in the UK and in Italy and I test for the presence of short run and long run asymmetries in price transmission using the nonlinear ARDL approach by Shin, Yu & Greenwood-Nimmo (2014), the Engle-Granger-Dickey-Fuller (EGDF) test and the bounds test by Pesaran, Shin & Smith (2001) which does not need the order of cointegration to be established before testing. I also account for possible effects of the exchange rate in price dynamics as suggested from works of Reilly & Witt (1998), Asplund, Eriksson & Friberg (2000) and Galeotti, Lanza & Manera (2003). Results indicate that both markets suffer from short run asymmetries but, while in the UK asymmetries vanish in the long run, they persist in Italy.

The plan of this chapter is as follows: section two provides a description of models and test used, section three describes the dataset, the price transmission mechanism in the oil market and the empirical exercise, section four concludes.

2.2 Models

I first start the analysis using a simple symmetric static OLS regression with and without trend :

$$r_t = \alpha + \beta h_t + u_t \tag{2.1}$$

$$r_t = \alpha + \delta t + \beta h_t + u_t \tag{2.2}$$

where r_t is the retail price, $h_t = w_t + f_t$ is the wholesale price measured in the local currency, w_t is the gasoline wholesale price in dollars and f_t is the foreign exchange rate.

I perform a residual based Engle-Granger-Dickey-Fuller (EGDF) test following the approach by Engle & Granger (1987) to assess cointegration and then I test for complete pass-through using a Wald test for the null $\beta = 1$.

Then, I use a symmetric dynamic Autoregressive Distributed Lag (ARDL) regression with and without trend

$$\Delta r_t = \alpha + \rho r_{t-1} + \theta h_{t-1} + \sum_{j=1}^{p-1} \phi_j \Delta r_{t-j} + \sum_{j=0}^{q} \pi_j \Delta h_{t-j} + \epsilon_t \tag{2.3}$$

$$\Delta r_t = \alpha + \delta t + \rho r_{t-1} + \theta h_{t-1} + \sum_{j=1}^{p-1} \phi_j \Delta r_{t-j} + \sum_{j=0}^{q} \pi_j \Delta h_{t-j} + \epsilon_t \tag{2.4}$$

where Δ indicates the first difference operator and p and q the lag length. The symmetric long run parameter is given by $\beta = \frac{\theta}{\rho}$. I perform a Pesaran, Shin & Smith (2001)[PSS] bounds test for cointegration which allows to achieve valid inference in the presence of both stationary and non-stationary variables. Also in this case, I perform a test of complete pass-through in the long run using a Wald test for the null hypothesis $\beta = 1$.

To analyse the asymmetric hypothesis of ‘Rocket and Feathers’, it is necessary to decompose the wholesale price in positive and negative observations. I first use an asymmetric cointegrating static OLS regression with and without trend as in the symmetric case before

$$r_t = \alpha + \beta^+ h_t^+ + \beta^- h_t^- + u_t \tag{2.5}$$

$$r_t = \alpha + \delta t + \beta^+ h_t^+ + \beta^- h_t^- + u_t \tag{2.6}$$

in which the wholesale price in local currency h_t is decomposed as

$$h_t = h_0 + h_t^+ + h_t^- \tag{2.7}$$

where h_0 is the initial value and where $h_t^+ = \sum_{j=1}^{t} \Delta h_j^+ = \sum_{j=1}^{t} \max(\Delta h_t, 0)$ and $h_t^- = \sum_{j=1}^{t} \Delta h_j^- = \sum_{j=1}^{t} \min(\Delta h_t, 0)$ are partial sum processes of positive and negative changes in h_t. Following Shin et al. (2014), I assume a single known threshold value of zero to ensure that the model retains a clear economic interpretation. Doing so, the observations are divided in about 45:55 in favour of the negative regime. This model allows for asymmetry as it has different parameters for positive and negative values of h_t.

Repeating the procedure for the symmetric case, I perform an EGDF test for cointegration and a Wald test for symmetry, $H_0 : \beta^+ = \beta^-$ and one for complete pass-through $H_0 : \beta^+ = 1; \beta^- = 1$.

Following Shin, Yu & Greenwood-Nimmo (2014), I employ the non-linear ARDL extension of (2.5) and (2.6), with and without trend respectively, to analyse the short run adjustment asymmetry and the long run reaction asymmetry:

$$\Delta r_t = \alpha + \rho r_{t-1} + \theta^+ h_{t-1}^+ + \theta^- h_{t-1}^- + \sum_{j=1}^{p-1} \phi_j \Delta r_{t-j} + \sum_{j=0}^{q} (\pi_j^+ \Delta h_{t-j}^+ + \pi_j^- \Delta h_{t-j}^-) + \epsilon_t \tag{2.8}$$

$$\Delta r_t = \alpha + \delta t + \rho r_{t-1} + \theta^+ h^+_{t-1} + \theta^- h^-_{t-1} + \sum_{j=1}^{p-1} \phi_j \Delta r_{t-j} + \sum_{j=0}^{q} (\pi^+_j \Delta h^+_{t-j} + \pi^-_j \Delta h^-_{t-j}) + \epsilon_t \tag{2.9}$$

This model allows for both short run and long run asymmetry; $\beta^+ = -\theta^+/\rho^+$ and $\beta^- = -\theta^-/\rho^-$ are the asymmetric long run parameters.

The PSS bounds test for cointegration is robust to non-stationary variables and it can be employed in this case. Long run asymmetry can be tested with a Wald test for the null $\beta^+ = \beta^-$ while complete pass-through can be tested setting the null $\beta^+ = 1; \beta^- = 1$. short run impact asymmetry can be tested on parameters π^+_j and π^-_j directly or, for additive asymmetry on the sum across lags of them with a Wald test.

As a robustness check, I then repeat all the procedure using f_t, the exchange rate and w_t, the gasoline wholesale price, as separate regressors in all the above models.

2.3 Data and Empirical results

I consider 111 monthly observations over the period January 1999 to March 2008 for gasoline wholesale and retail prices in the UK and in Italy. As these two countries pay for crude oil in US dollars, it is necessary to consider the exchange rate for Dollar, Euro and Pound sterling either as a regressor or to convert the wholesale price in local currency.

The gasoline wholesale price is the Amsterdam-Rotterdam-Antwerp (ARA) 50 parts per million conventional gasoline regular spot in dollars per gallon obtained from the International Energy Agency.

The gasoline retail prices refer to unleaded gasoline price charged by retail stations to customers excluding taxes in Euro per litre for Italy and Pound sterling per litre for the UK. All variables are logged.

Figure 2.1: Plot of wholesale log-prices (red line) and retail log-prices (blue line)

(a) UK (b) Italy

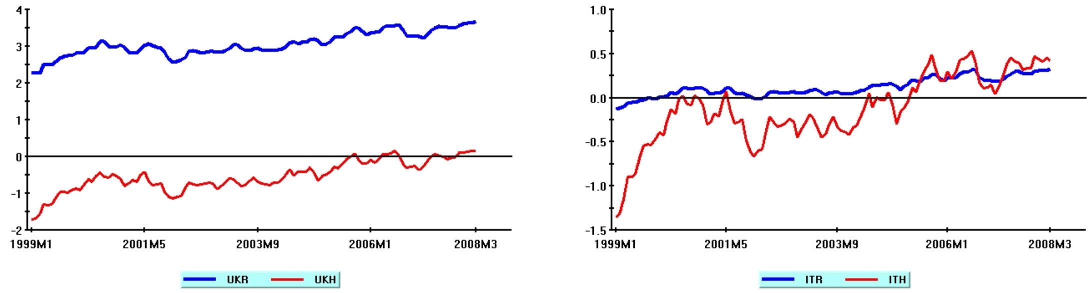
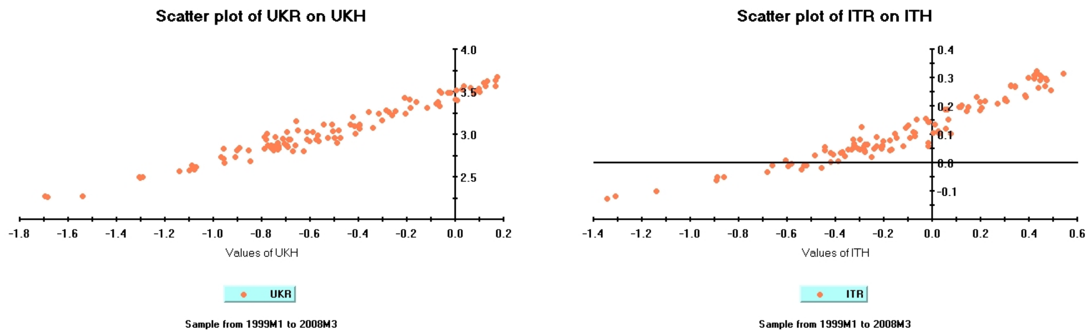


Figure 2.2: Scatterplot of retail log-prices against wholesale log-prices

(a) UK (b) Italy



At a first analysis of the plot of the UK and Italian series in figure 2.1, it seems there is an upward comovement that suggests a cointegration relation. Moreover, in the Italian case, from 2005, wholesale prices exceeded retail prices in some occasions. Looking at scatter plots in figure 2.2, it looks like this relation is linear and the xplot of figure 2.3 shows a zigzag pattern so no clear known distribution is recognisable. Plotting the prices first differences in figure 2.4, I can see that difference wholesale prices show higher volatility than difference retail prices especially in the Italian

Figure 2.3: Xplot of retail log-prices against wholesale log-prices

(a) UK (b) Italy

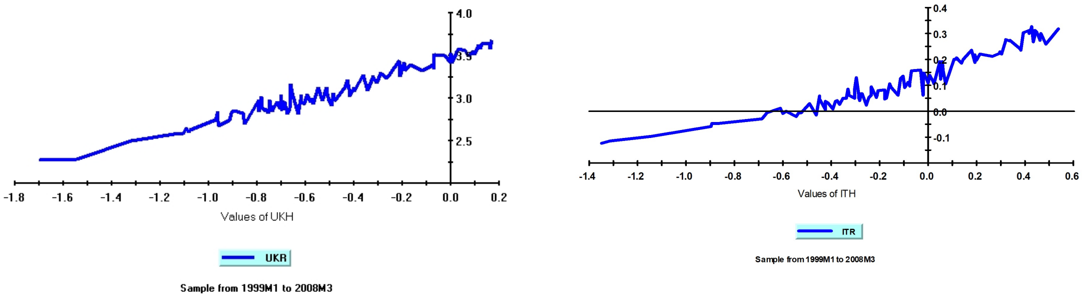
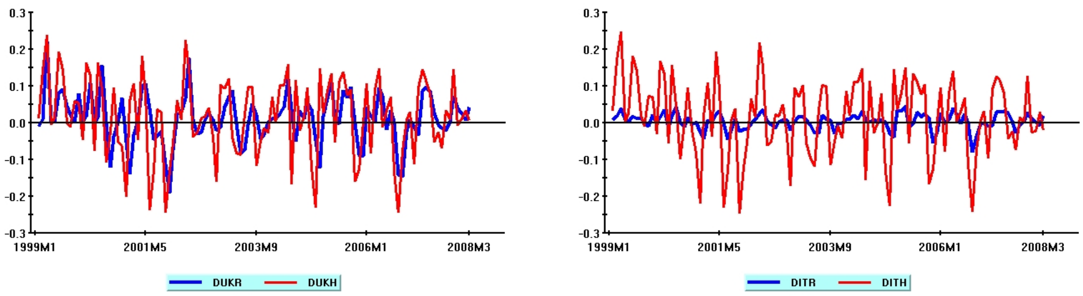


Figure 2.4: Plot of first difference retail log-prices (blue line) against wholesale log-prices (red line)

(a) UK (b) Italy



case. Scatter plots of variables first differences in figure 2.5 show that the relation among prices can be linear but in the first quadrant it seems stronger and in the third quadrant weaker and this can suggest asymmetry.

Table 2.1 lists all the models considered and tests performed for both the British and Italian markets. For the ARDL specification, following the general-to-specific lag selection approach, I start from the maximum lag order of twelve and, using the regression results time after time obtained, I drop not significant regressors ending up using always only two lags.

Table 2.1: Models and Tests performed in the empirical exercise repeated for both the UK and Italy

Empirical Exercise Steps

1) Symmetric Static OLS Regression
- a. EGDF Test (Unit root test)
- b. Wald Test $H_0 : \beta = 1$ (Complete pass-through)

2) Symmetric Dynamic ARDL
- a. PSS Bounds F-Test for cointegration
- b. Wald Test $H_0 : \beta = 1$ (Complete pass-through)

3) Asymmetric Static OLS Regression
- a. EGDF Test (Unit root test)
- b. Wald Test $H_0 : \beta^+ = 1$ and $\beta^- = 1$ (Complete pass-through)
- c. Wald Test $H_0 : \beta^+ = \beta^-$ (Symmetry)

4) Asymmetric ARDL Regression
- a. PSS Bounds F-Test for cointegration
- b. Wald Test $H_0 : \beta^+ = 1$ and $\beta^- = 1$ (Complete pass-through)
- c. Wald Test $H_0 : \beta^+ = \beta^-$ (Symmetry)
- d. Wald Test $H_0 : \pi_0^+ = \pi_0^-$ and $H_0 : \pi_1^+ = \pi_1^-$ (Impact Symmetry)
- e. Wald Test $H_0 : \pi_0^+ + \pi_1^+ = \pi_0^- + \pi_1^-$ (Additive Symmetry)

Table 2.2: Static OLS regression estimation and tests results for the UK and Italy. * indicates significance at 5%.

Symmetric Static OLS regression

Coefficient	**UK Estimate**	**IT Estimate**
Constant	3.332	0.073
Trend	0.002	0.001
β	0.667	0.186

Test H_0	**UK Statistic**	**IT Statistic**
$\beta = 1$	125.54*	7690.90*

Asymmetric Static OLS Regression

Coefficient	**UK Estimate**	**IT Estimate**
Constant	2.189	-0.185
β^+	0.694	0.207
β^-	0.663	0.182

Test H_0	**UK Statistic**	**IT Statistic**
$\beta^+ = 1$ and $\beta^- = 1$	212.07*	16309.40*
$\beta^+ = \beta^-$	13.45*	104.13*

Table 2.3: Dynamic ARDL regression estimation and tests results for the UK and Italy. * indicates significance at 5%.

Symmetric Dynamic ARDL

Coefficient	**UK Estimate**	**IT Estimate**
Constant	1.059	0.018
Trend	N/A	0.012
ρ	-0.303	-0.216
θ	0.244	0.049
$\sum \phi$	0.178	0.046
$\sum \pi$	0.528	0.209

Test H_0	**UK Statistic**	**IT Statistic**
$\rho = \theta = 0$	12.10*	9.35*
$\beta = 1$	52.54*	1053.60*

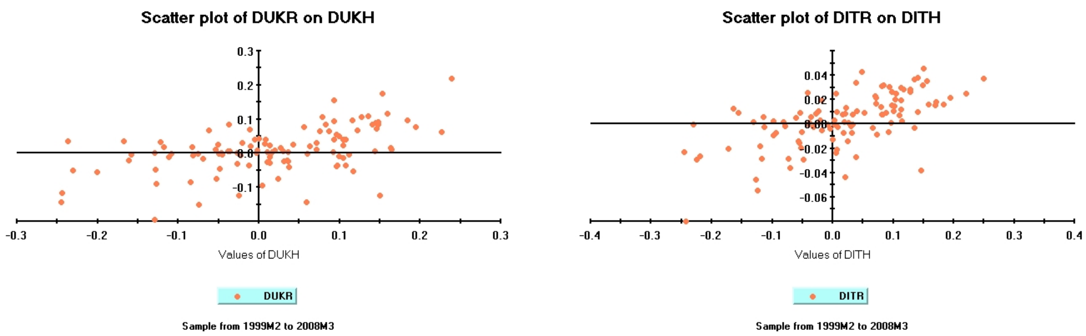
Asymmetric Dynamic ARDL

Coefficient	**UK Estimate**	**IT Estimate**
Constant	0.632	-0.043
ρ	-0.297	-0.207
θ^+	0.225	0.050
θ^-	0.219	0.045
ϕ	0.181	0.034
$\sum \pi^+$	0.572	0.185
$\sum \pi^-$	0.484	0.233

Test H_0	**UK Statistic**	**IT Statistic**
$\rho = \theta^+ = \theta^- = 0$	7.64*	7.10*
$\beta^+ = \beta-$	1.90	16.18*
$\beta^+ = 1$ and $\beta^- = 1$	52.06*	2002.50*
$\sum \pi^+ = \sum \pi^-$	0.47	1.17
$\pi_0^+ = \pi_0^-$	7.47*	1.05
$\pi_1^+ = \pi_1^-$	3.35*	6.43*

Figure 2.5: Scatterplot of first difference retail log-prices against wholesale log-prices

(a) UK (b) Italy



2.3.1 The UK

First, running the symmetric OLS regressions both with intercept and time trend and intercept only (equations (2.1) and (2.2)), I find that the time trend is significative so the model with time trend is preferable; also for the latter, the $\bar{R}^2$ is higher, but presence of serial correlation in residuals and heteroscedasticity are detected by the Durbin-Watson test and test of unconditional homoscedasticity assumption respectively. I perform the EGDF residual based test for cointegration and, for both models, the AIC suggests using an Augmented Dickey Fuller test with 1 lag. Comparing its t-test value with the DF critical value, I can reject the null of no cointegration. The Wald test for complete pass-through ($H_0 : \beta = 1$) strongly rejects the null, so it appears that variations of wholesale prices are not passed entirely to retail prices.

Then, I use the symmetric ARDL regression of equations (2.3) and (2.4), following the general-to-specific lag selection rule, from the regression results obtained using ARDL(12,12) with time trend and intercept, I find that coefficients with lag more than 2 and time trend are statistically insignificant so I end up using the more parsimonious ARDL(2,2) model without time trend.

In this model, all regressors appear to be highly statistically significant and no serial correlation is detected; Ramsey (1969) test for functional form does not reject the null of linear misspecification and the null hypothesis of homoscedasticity cannot be rejected with a level of significance less than 5%. Performing the PSS Bounds test case III for the long run level cointegrating relation ($H_0 : \rho = \theta = 0$), I strongly reject the null of no level relation so I can estimate the long run multiplier as $\beta = -\theta/\rho$ and testing for $H_0 : \beta = 1$ (complete pass-through) I reject the null, even if there is no complete pass-through, the estimate of β is 0.805, so retail prices appear to be quite responsive to variations in wholesale prices. Almost identical results are also obtained using an ARDL (12,12).

Turning to asymmetric models, first I use the static OLS models (2.5) and (2.6). Also in this case, time trend appears not to be statistically significative so I use the model (2.6) without a time trend, in this regression all coefficients are significative but Godfrey test for serial correlation suggest serial correlation and the presence of homoscedaticity is also detected.

I perform the EGDF residual based test for cointegration; the AIC and other criteria again suggest an ADF(1) and I can reject the null hypothesis of no cointegration. The estimated values of β^+ and β^- are, respectively 0.69421 and 0.66318 so it seems to be only a subtle difference between the two but performing a Wald test for symmetry ($H_0 : \beta^+ = \beta^-$), I can strongly reject the null so data suggest asymmetry in effects of positive and negative wholesale price changes. Moreover, estimations suggest that positive variations in wholesale prices have slightly more effect than negative one on retail prices.

I also perform a Wald for complete pass-through ($H_0 : \beta^+ = 1; \beta^- = 1$) and the null is rejected confirming changes in wholesale prices are not entirely passed to retail prices.

Then, I start using asymmetric ARDL of equations (2.8) and (2.9), following the general-to-specific lag selection approach, I end up using an ARDL(2,2,2) without time trend as other lags are not significative. Godfrey test suggests no serial correlation but heteroscedasticity appears.

I perform the PSS Bounds test case III for cointegration ($H_0 : \rho = \theta^+ = \theta^- = 0$) because it is preferable to the EGDF as the latter suffers of small power in finite sample, the value of the F-test falls outside the bounds for size 1% so I reject the null hypothesis of no cointegration. I calculate the long run multipliers and I test them for symmetry ($H_0 : \beta^+ = \beta^-$) using a Wald test and I cannot reject the null of symmetry in the long run, which means that data suggest positive and negative variations of crude oil affects the retail price in the same way in the long run. I also perform a complete pass-through Wald test on long run multipliers ($H_0 : \beta^+ = 1; \beta^- = 1$) and the null hypothesis of complete pass-through is strongly rejected.

At this stage, I perform some tests to assess the presence of asymmetry in the short run: the first regards impact symmetry ($H_0 : \pi_0^+ = \pi_0^-$ and $H_0 : \pi_1^+ = \pi_1^-$) and I can reject the null hypothesis, as impact asymmetry restrictions are likely to be excessively tightening, I perform a test for additive symmetry ($H_0 : \pi_0^+ + \pi_1^+ = \pi_0^- + \pi_1^-$) but in this case I cannot reject the null hypothesis so data support impact asymmetry but not additive asymmetry.

Performing regressions adding the logarithm of exchange rate as a regressor, I mainly find the same results but estimates for exchange rates seems to be not much significant in ARDL cases. All tests show symmetric impact of forex variations both in the long run and the short run. The only contrasting result is that using asymmetric OLS (models (2.1) and (2.6)) I cannot reject the null of symmetry for wholesale prices.

2.3.2 Italy

First, running symmetric OLS regressions both with intercept and time trend and only with intercept (equations (2.1) and (2.2)), I find that time trend is significative so the model with time trend is preferable; also the $\bar{R}^2$ is higher in this case, but presence of serial correlation is detected by tests. Moreover, Ramsey's test for functional form suggests that the linear form may not be right. I perform the EGDF residual based test for cointegration and for both models, different criterion suggest different lag order and the test fails to reject the null of no cointegration starting from the first lag but the EGDF test is known to have too small power in finite sample and it may fail to reject the null of no cointegration even if there is cointegration. The Wald test for complete pass-though ($H_0 : \beta = 1$) strongly rejects the null, so variations of wholesale prices are not passed entirely to retail prices.

Then, I use the symmetric ARDL regression of equations (2.3) and (2.4), following the general-to-specific lag selection rule, from the regression results obtained using ARDL(12,12) with time trend and intercept, I find that coefficients with lag more than 2 are statistically insignificant so I use only an ARDL(2,2) with time trend differently from the UK case. In this model, all regressors appear to be highly statistically significant except for the difference of the retail prices, and no serial correlation is detected using 10% size, test for functional form does not reject the null of linear specification and the null hypothesis of homoscedasticity is rejected.

Performing the PSS Bound test for long run level relation case V ($H_0 : \rho = \theta = 0$), I reject the null of no level relation and the long run multiplier is given by $\beta = -\theta/\rho$; I test for $\beta = 1$ (complete pass-through) and I reject the null, the estimate of β is 0.227, so retail prices are less responsive to wholesale prices than in the UK case. Almost identical results are also obtained using an ARDL (12,12).

Turning to asymmetric models, first I use static OLS on models (2.5) and (2.6). In this case, the time trend appears not to be statistically significative so I use model (2.5), in this regression all coefficients are significative but Godfrey test for serial correlation suggests serial correlation. I perform the EGDF residual based test for cointegration; different criteria suggest different lags, I can only reject the null of no cointegration with ADF(1) but the EGDF may have bad power performances in small samples. The estimated values of β^+ and β^- are, respectively 0.207 and 0.182 so it seems the two are only slightly different but performing a Wald test for symmetry in the long run ($H_0 : \beta^+ = \beta^-$) I reject the null hypothesis so data suggest asymmetry in effects of positive and negative wholesale price changes.

I also perform a Wald test for complete pass-through ($H_0 : \beta^+ = 1; \beta^- = 1$) and the null is rejected so changes in wholesale prices are not entirely passed to retail prices as in the UK.

Using asymmetric ARDL of equations (2.8) and (2.9), following the general-to-specific approach, I end up using an ARDL(2,2,2) without time trend as other lags are not significative but the coefficient for the differences in retail prices is still not very much significative. Godfrey test for serial correlation suggests no serial correlation but heteroscedasticity appears to be present.

I perform the PSS Bounds test for cointegration case III because it is preferable to the EGDF as the latter has too small power in finite sample. The value of the F-test for the null of $\rho = \theta^+ = \theta^- = 0$ falls outside the bounds for size 1% so I reject the null of no cointegration. I calculate the long run multipliers and I test them for symmetry using a Wald test and I reject the null hypothesis of symmetry in the long run ($H_0 : \beta^+ = \beta^-$) which means that data suggest positive and negative variations of crude oil affect the retail price in a different way in the long run differently from the UK case. I also perform a complete pass-through Wald test on long run multipliers

$(H_0 : \beta^+ = 1; \beta^- = 1)$ and it strongly rejects the null of complete pass-through.
Now, I perform some test to assess the presence of asymmetry in the short run; the first one regards the impact symmetry $(H_0 : \pi_0^+ = \pi_0^-)$ and I can reject the null. As the impact asymmetry restrictions are likely to be excessively restrictive, I perform a test for sum symmetry $(H_0 : \pi_0^+ + \pi_1^+ = \pi_0^- + \pi_1^-)$ and I can reject the null of symmetry even in this case. Using the exchange rate as regressor I obtain mainly the same results but using asymmetric ARDL model I cannot reject the null of both long run and short run symmetry regarding wholesale price and foreign exchange.

2.4 Conclusions

In this chapter, I check how variations in gasoline wholesale prices influence retail prices in the UK and Italy using both a standard OLS regression and an ARDL model. The latter can be estimated using OLS in only one step avoiding uncertainty arising from a previous estimation step as in the Engle & Granger (1987) framework. The hypothesis of no cointegration can be tested using the PSS Bounds test which is valid in the presence of both stationary and non stationary variables. Long run symmetry and short run impact and additive symmetry can be tested using a Wald test.
The effect of the exchange rate is kept into account converting the wholesale prices in local currency and then repeating the empirical analysis using the exchange rate as an additional regressor instead. Both methods give similar results.
Analysing data for the UK, it emerges that there is asymmetry in the short run but in the long run no asymmetry is detected. Retail prices slowly reflect changes in wholesale prices both positive and negative. Differently, in Italy, I detect asym-

metries both in the long and short run corroborating the general perception that negative variations of wholesale prices are not passed to consumers.

In both cases, it is plausible to detect asymmetry, at least in the short run, due to the market power of firms in this sector and the inelastic demand. In Italy, fixed taxes are an important component of the pump price, as these are not subject to fluctuations, this can cause prices to not completely reflect changes in wholesale prices. At the same time, producers may not be willing to decrease pump prices fearing the government will add more taxes dissolving the advance for final consumers and frustrating the price decrease. For future work, more reliable results might be obtained using daily data instead of monthly as variations can occur between two monthly observations and these movements can be captured using more frequent observations.

Chapter 3

Tests for Equal Predictive Accuracy in the Presence of Heteroscedasticity and Serial Correlation

Abstract

This chapter presents three of the most popular tests for model-free forecast evaluation and three other tests when forecasts are model-based. It then focuses on the Diebold and Mariano test as it is the one based on less restrictive assumptions on forecast errors and it is widely used also for comparing model-based forecasts. In the presence of serial correlation and heteroscedasticity, this test suffers from size distortion which can be alleviated using fixed-smoothing asymptotics for critical values.

Montecarlo results using fixed b and fixed m approaches are promising. The same Montecarlo exercise is repeated using the Giacomini and White test for model-based forecasts in small samples obtaining good results. An empirical example about the Dollar/Euro exchange rate comparing random walk forecasts and three months forward rates is also presented. Results indicate random walk forecasts predict better the trend of the exchange rate considered.

3.1 Introduction

Good forecasts are pivotal in decision making, economic and financial research and policy making; since there are different acceptable methods to forecast a time series, there is the need for methods to compare predictive accuracy to discern between good and bad forecasts.

Forecast evaluation tests can be model-based or model-free: model-based tests assume forecasts from a parametric econometric model estimated from a sample, while model-free tests assume that the only thing available is a set of forecasts and actual values of the variable of interest.

To assess the quality of forecasts, it is possible to compare an error measure like the Mean Square Error but as Diebold (2015) notices, the fact that the MSE for a forecast is smaller than for another may depend on a specific sample realization and not be true in population. Model-free tests that compare the forecast of two alternative time series methods include Morgan-Granger-Newbold (1977) (Morgan, 1939; Granger & Newbold, 1986), and an extension by Meese & Rogoff (1988) to account for serially correlated forecast errors, and one of the most used test in this field: the Diebold and Mariano test (1995). These tests are based on a loss function associated with the forecast error of each forecast, testing the null hypothesis of zero expected loss differential of the two competing forecasts.

In small samples, these tests suffer from size distortion as noted by Clark (1999) and others. This issue can be alleviated using a heteroscedasticity autocorrelation robust approach for the long run variance estimator. In this light, Coroneo & Iacone (2015) and Harvey, Leybourne & Whitehouse (2016) perform a study of the Diebold and Mariano test combined with fxed b asymptotics by Kiefer & Vogelsang (2005) and fixed m asymptotics by Hualde & Iacone (2015) in the case of serial correlation only and these prove to be good in delivering correctly sized tests.

When considering model-based forecasts, West (1996) develops a test for non nested models based on Diebold & Mariano (1995) while Clark & McCracken (2001), Clark & McCracken (2005) and McCracken (2007) introduce a set of asymptotics for comparison of forecasts from two different nested models at population level. Giacomini & White (2006) instead, deal with equal forecast accuracy in a finite sample setting when estimation follows a rolling scheme with a finite observation window. Their test is equivalent to a Diebold and Mariano test. In addition, West (1996) reports that, in some cases, the original Diebold and Mariano approach is still valid when forecasts come from a specific model. For these reasons, the focus of this work is mainly on the Diebold and Mariano test.

In this chapter, after a rapid overview of major tests for equal forecast accuracy in model-free and model-based settings, I perform a Montecarlo study about size and power of the Diebold and Mariano test statistic in case of heteroscedasticity comparing fixed-smoothing asymptotics to standard asymptotics obtaining promising results. The same exercise is also performed on the Giacomini & White (2006) test. I then use the Diebold and Mariano test and fixed-smoothing asymptotics to assess whether a random walk forecast and a three-months forward rate have equal forecast accuracy about the Dollar/Euro exchange rate in a similar way the original Diebold & Mariano (1995) paper was doing with the Dollar/Dutch guilder exchange rate.

The remainder of this chapter is organised as follows: section two and three give a broad overview of equal forecast accuracy tests in model-free and model-based settings respectively, in the fourth section, Weighted Covariance Estimators and Weighted Periodogram Estimators for the long run variance are discussed along with fixed b and fixed m approaches.

The fifth section provides Montecarlo results for size and power of Diebold and Mariano and Unconditional Giacomini and White tests comparing standard asymptotics to fixed-smoothing asymptotics. In section six, an empirical application of Diebold and Mariano test checking the predictive accuracy of forward rates against a random walk using all the different asymptotics is discussed and section seven concludes.

3.2 Model-free Tests for Equal Forecast Accuracy

Consider a time series of actual values of the variable of interest y_t observed at times $t = 1, 2, 3, ..., T$. For this time series two competing forecasts are available: $\hat{y}_{t1}$ and $\hat{y}_{t2}$ with $t = 1, 2, 3, ..., T$. Forecast errors are defined as $e_{ti} = \hat{y}_{ti} - y_t$ for $i = 1, 2$. Tests for forecast accuracy consider the expected loss associated with each forecast. The loss is a function of the error $l(y_t; \hat{y}_{ti}) = l(\hat{y}_{ti} - y_t) = l(e_{ti})$. Usually this loss function is the square or the absolute value of e_{ti} but other specifications to include functional dependence or asymmetry are possible, like for example, the one suggested in Clements & Hendry (1993), Pesaran & Timmermann (1994), Christoffersen & Diebold (1996), Stock & Watson (1999) and Granger & Pesaran (2004) among others.

The loss differential for two forecasts is defined as

$$d_t = l(e_{t1}) - l(e_{t2}); \tag{3.1}$$

two forecasts have equal predictive accuracy if the expectation of the loss differential is zero, so the null hypothesis to test is $H_0 : E[d_t] = 0$, no difference in forecast accuracy, against the alternative $H_1 : E[d_t] \neq 0$.

3.2.1 Morgan-Granger-Newbold Test

Morgan, Granger and Newbold (Morgan, 1939; Granger & Newbold, 1986) uses a quadratic loss function and assumes the forecast errors are:

i zero mean,

ii Gaussian,

iii serially uncorrelated,

This test is based on a orthogonalising transformation which makes it robust to contemporaneous autocorrelation in the forecast errors.
Using these forecast error vectors:

$$x_t = e_{t1} + e_{t2} \quad \text{and} \quad y_t = e_{t1} - e_{t2} \tag{3.2}$$

we have $E[x_t y_t] = E[(e_{t1} + e_{t2})(e_{t1} - e_{t2})] = E[e_{t1}^2 - e_{t2}^2] = \sigma_1^2 - \sigma_2^2$ where σ_1^2 and σ_2^2 are the variances of the forecasts errors. Given the loss function is quadratic, the null hypotheses $H_0 : E[d_t] = 0$ is equivalent to $E[x_t y_t] = 0$ and the Morgan-Granger-Newbold test of zero correlation between x_t and y_t, using results from Hogg & Craig (1970), is

$$MGN = \frac{\hat{\rho}_{x_t y_t}}{\sqrt{(n-1)^{-1}(\hat{\rho}^2_{x_t y_t})}} \quad \sim \quad t_{T-1} \quad \text{under} \quad H_0 \tag{3.3}$$

where $\hat{\rho}_{x_t y_t} = \frac{x'y}{\sqrt{(x'x)(y'y)}}$ is the estimated contemporaneous correlation.
This distribution result is valid for any sample size and ρ under assumption ii. Assumption iii implies the test is only valid for 1-step-ahead forecast because, if the horizon is bigger than one, autocorrelation may be a MA($h-1$) process.
The test is valid as a test of equality of forecast accuracy only under squared error loss; moreover, simulations in Harvey (1997) show that it is oversized for non-normal errors which are very common in practice.

3.2.2 Meese-Rogoff Test

To overcome the problem of autocorrelation, Meese & Rogoff (1988) relax the assumptions and allow the forecast errors to be serially correlated. Using a quadratic loss function under assumption i and ii from previous section, they show that, as $T \to \infty$,

$$\sqrt{T}\hat{\gamma}_{xy} \xrightarrow{d} N(0, \Sigma) \tag{3.4}$$

where $\hat{\gamma}_{xy} = \frac{x'y}{T}$ and $\Sigma = \sum_{\tau=-\infty}^{\infty}[\gamma_{xx}(\tau)\gamma_{yy}(\tau) + \gamma_{xy}(\tau)\gamma_{yx}(\tau)]$ and $\gamma_{xx}(\tau) = cov(x_t; x_{t-\tau})$, $\gamma_{yy}(\tau) = cov(y_t; y_{t-\tau})$, $\gamma_{yx}(\tau) = cov(y_t; x_{t-\tau})$, $\gamma_{xy}(\tau) = cov(x_t; y_{t-\tau})$.
Here and after, $\xrightarrow{d}$ indicates convergence in distributions as $T \to \infty$.
These covariances can be estimated using their sample analogues and a consistent

estimator of Σ is

$$\hat{\Sigma} = \sum_{\tau=-M}^{M} \left[1 - \frac{|\tau|}{T}\right] [\hat{\gamma}_{xy}(\tau)\hat{\gamma}_{yx}(\tau) + \hat{\gamma}_{xx}(\tau)\hat{\gamma}_{yy}(\tau)] \tag{3.5}$$

where M is the truncation lag which increases with the sample size T but at a slower rate.

Their test statistic based on x and y is

$$MG = \frac{\hat{\gamma}_{xy}}{\sqrt{\frac{\hat{\Sigma}}{T}}} \quad \xrightarrow{d} \quad N(0,1) \quad \text{under} \quad H_0 \tag{3.6}$$

Under assumption i, ii and iii, MR coincides asymptotically with MGN.

3.2.3 Diebold and Mariano Test

Diebold & Mariano (1995) propose a model-free test for the null hypothesis of no difference in the accuracy of two forecasts applicable to non-quadratic loss functions and multi-period forecasts. All assumptions on forecast errors and loss function are relaxed making this test widely applicable.

The Diebold and Mariano test is based on the sample mean of the observed loss differential series

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t \tag{3.7}$$

The null hypotheses of equal forecast accuracy is made on the population mean of the loss differential series as $H_0 : \mu = 0$ where $\mu = E[d_t]$.

If the series of loss differentials is covariance stationary and short memory, standard results of the Central Limit Theorem can be used to deduce the asymptotic distribution of the sample mean loss differential to get

$$\sqrt{T}(\bar{d} - \mu) \xrightarrow{d} N(0, \sigma^2)$$

The parameter σ^2 is known as long run variance.

Using standard results,

$$\sigma^2 = 2\pi f_d(0) = 2\pi \tfrac{1}{T} \textstyle\sum_{\tau=-\infty}^{\infty} \gamma_\tau$$

where $f_d(\lambda) = \frac{1}{T}\sum_{\tau=-\infty}^{\infty} \gamma_\tau e^{-i\tau\lambda}$ is the spectral density of the loss differential, in this case, evaluated at frequency $\lambda = 0$ and $\gamma_\tau = E[(d_t - \mu)(d_{t-\tau} - \mu)]$ is the autocovariance of the loss differential at displacement τ.

Because in large samples the sample mean of the loss differential is approximately normally distributed with mean μ and variance σ^2, the large sample test statistic for the null hypothesis of equal forecast accuracy is

$$\sqrt{T}\frac{\bar{d} - \mu}{\sigma} \quad \xrightarrow{d} \quad N(0,1) \quad \text{under} \quad H_0 \tag{3.8}$$

This test statistic cannot be calculated because σ is unknown. Diebold & Mariano (1995) use the following estimate

$$\hat{\sigma}^2 = 2\pi \hat{f}_d(0) = \sum_{\tau=-M(T)}^{M(T)} k\left(\frac{\tau}{M(T)}\right)\hat{\gamma}_\tau \tag{3.9}$$

where $\hat{\gamma}_\tau = \frac{1}{T}\sum_{t=|\tau|}^{T}(d_t - \bar{d})(d_{t-|\tau|} - \bar{d})$, $M(T)$ is the lag truncation and $k(.)$ is the kernel function.

Then, they use a Montecarlo simulation to compare these three tests: results show that, in case of Gaussian forecast errors, MGN is correctly sized if there is no serial

correlation in errors; contemporaneous and serial correlation does not affect MR in large samples but serial correlation in small samples makes it slightly oversized.
The behaviour of Diebold and Mariano test statistic is similar to that of MR: robust to contemporaneous and serial correlation in large samples and oversized in small samples.
When forecast errors are non-Gaussian, MGN and MR suffer from severe size distortion in both large and small samples. Diebold and Mariano test, on the other hand, still has approximately correct size for all sample sizes except when they are very small.

3.3 Model-based Tests for Equal Forecast Accuracy

This section is dedicated to present tests available for cases in which two alternative models are used to forecast the variable of interest y_t. In this setting, competing forecasts are depending on model estimated parameters which are updated as time passes and this introduces a new source of variation in forecasts. West (1996) focuses on developing methods for testing the accuracy of forecasts at unknown population values of parameters, for instance an infinite sample of data is available to estimate model parameters while, other approaches, like the one presented in Giacomini & White (2006), are concerned with testing for equal forecast accuracy in finite samples.

3.3.1 Non-Nested Models

Let $L_{t+h|t}(\beta)$ be a vector of losses whose individual elements are denoted by $l_{t+h|t,i}(\beta)$; $l_{t+h|t}(\hat{\beta}_t)$ is the observed loss on the parameter estimate $\hat{\beta}_t$. β^* is the limiting value of $\hat{\beta}_t$ as the sample size gets very large and $l_{t+h|t}(\beta^*)$ is the loss function evaluated at this pseudo-true value, while $F = E[\partial l_{t+h|t}(\beta)/\partial\beta]$ is the derivative of each individual element of the loss vector evaluated at β^*. Also, η_{t+h} is the orthogonality condition used to estimate the model parameters so that $\eta_{t+h}(\beta^*)$ is a zero-mean process. Under the following assumptions:

i $l_{t+h|t}(\beta)$ is measurable and twice differentiable function in some open neighbourhood around β^*;

ii the estimate $\hat{\beta}_t$ can be written as a moment-type estimator of full rank: $\hat{\beta}_t - \beta^* = B(t)H(t)$, where $B(t)$ is $k \times q$, $H(t)$ is $q \times 1$ with $B(t) \to_{as} B$, for B of rank k, $H(t) = (t-h)^{-1}\sum_{s=1}^{t-h}\eta_{s+h}(\beta^*)$ where $\eta_{s+h}(\beta^*)$ is an orthogonality condition and $E[\eta_s(\beta^*)] = 0$;

iii $|\partial^2 l_{t+h|t}(\beta)/\partial\beta\partial\beta'|$ is bounded;

iv $\partial l_{t+h|t}(\beta)/\partial\beta$ evaluated at β^* satisfies certain mixing, moment and stationarity conditions. Moreover $S_y(0) = \sum_{j=-\infty}^{\infty} E[(l_{t+h|t}(\beta^*) - E[l_{t+h|t}(\beta^*)])\times$ $(l_{t-j+h|t-j}(\beta^*)' - E[l_{t-j+h|t-j}(\beta^*)'])]$ is assumed to positive definite;

v T_R is the number of observations used for estimation, T_P the number of observations used for evaluation, h is the forecast horizon, $T_P = T - T_R - h \to \infty$ and $T_P/T_R \to \pi$. Defining: $\Pi = \begin{cases} 0, & \text{for } \pi = 0 \\ 1 - \ln(1+\pi)/\pi, & \text{for } 0 < \pi < \infty \\ 1, & \text{for } \pi = \infty \end{cases}$

West (1996, Theorem 4.1) establishes that

$$T_P^{-1/2}\left(\sum_{t=T_R}^{T-h} l_{t+h|t}(\hat{\beta}_t) - E[l_{t+h|t}(\beta^*)]\right) \to_d N(0,\sigma^2) \tag{3.10}$$

where

$$\sigma^2 = S_y(0) + \Pi(FBS'_{yh}(0) + S_{yh}(0)B'F') + 2\Pi FV_\beta F' \tag{3.11}$$

$S_{yh}(0) = \sum_{j=-\infty}^{\infty} E[\eta_t\eta'_{t-j}]$ and V_β is the limiting variance-covariance matrix of the estimators.

$S_y(0)$ is the long run variance of the loss under known model parameters, $2\Pi FV_\beta F'$ reflects how parameter estimation affects the variance of the loss and $\Pi(FBS'_{yh}(0)+S_{yh}(0)B'F')$ reflects the covariance between the long run variance of the loss and the effect of parameters estimation on it. Since the second and third terms reflect the estimation error, they will depend on the underlying estimation scheme via Π. A consistent estimate for σ^2 can be obtained setting

$$\Pi = 1 - \left(\frac{T_P}{T_R}\right)^{-1} \ln\left(1 + \frac{T_P}{T_R}\right) \tag{3.12}$$

$$\hat{F} \equiv P^{-1}\sum_{t=T_R}^{T} \frac{\partial l_{t+h|t}}{\partial\beta}(\hat{\beta}_t) \to_p F \tag{3.13}$$

and $\hat{B} \equiv B(T) \to_p B$ by assumption ii.

This result establishes that asymptotically, inference about forecast accuracy can be carried out using a standard normal distributions provided that the correct estimate of σ^2 is used. In practice, t-tests for equal mean squared error can be performed as in Diebold & Mariano (1995) without correcting for the effect of parameters estimation.

3.3.2 Nested Models

When comparing two nested models under the null hypothesis of equal forecast accuracy, West (1996)'s approach is no longer valid as the long run variance σ^2 becomes zero. This is due to the fact that forecast errors from competing models are always the same under the null.

Clark & McCracken (2001), Clark & McCracken (2005) and McCracken (2007) develop a theory for testing the null of equal forecast accuracy at population level for forecasts obtained from two nested models.

Given forecasts $\hat{y}_t(\hat{\beta})$, forecast errors $e_t(\hat{\beta}) = y_t - \hat{y}_t(\hat{\beta})$ where y_t is the true value of the variable of interest, and losses $l(e_t(\hat{\beta}))$ for each model, under the following sufficient assumptions:

i $l_{t+h|t}(\beta)$ is measurable and twice differentiable in some open neighbourhood around β^*;

ii the estimate $\hat{\beta}_t$ can be written as a moment-type estimator of full rank: $\hat{\beta}_t - \beta^* = B(t)H(t)$, where $B(t)H(t)$ equals $((t-h)^{-1}\sum_{s=1}^{t-h} \partial^2 l(\beta)/\partial\beta\partial\beta')^{-1}((t-h)^{-1}\sum_{s=1}^{t-h} \partial l_{t+h|t}(\beta)/\partial\beta)$,

iii $|\partial^2 l_{t+h|t}(\beta)/\partial\beta\partial\beta'|$ is bounded and $\partial l_{t+h|t}(\beta)/\partial\beta$ evaluated at β^* satisfies certain mixing, moment and stationarity conditions;

iv For the nesting model, $U_t = [h_t', vech(q_t - E[q_t])']'$ is covariance stationary with $h_t = \partial l_{t+h|t}(\beta^*)/\partial\beta^*$ and $q_t = \partial^2 l(\beta^*)/\partial\beta^*\partial\beta^{*'}$; $E[h_t|h_{t-j}, q_{t-j}, j > 0] = 0$;

$e[h_t h_t'] = cE[q_t] < \infty$; $\lim_{T\to\infty} T^{-1}E[(\sum_{s=1}^{T}\tilde{U}_s)(\sum_{s=1}^{T}\tilde{U}_s)'] = \sigma^2 < \infty$ is positive definite.

v both T_R, the number of observations used for estimation, and T_P, the number of observations used for evaluation, become arbitrarily large at the same rate,

the test statistic for the null $H_0 : E[l_t(\beta_1^*)] = E[l_t(\beta_2^*)]$ against the one sided alternative $H_1 : E[l_t(\beta_1^*)] > E[l_t(\beta_2^*)]$ is

$$CM = \frac{\sum_{t=T_R}^{T} l_t(\hat{\beta}_1) - l_t(\hat{\beta}_2)}{\hat{\sigma}\sqrt{T_P}} \to_d \frac{\Gamma_1 - 1/2\Gamma_2}{\sqrt{\Gamma_2}} \tag{3.14}$$

where $\Gamma_1 = \int_\lambda^1 s^{-1}W'(s)dW(s)$, $\Gamma_2 = \int_\lambda^1 s^{-2}W'(s)W(s)dW(s)$, $\lambda = (1+\pi)^{-1}$ and $W(s)$ is a standard Brownian motion.

Hansen & Timmermann (2012) show that the asymptotic distribution can be simplified to $\sqrt{1-\rho}(Z_1^2 - Z_2^2) - \log\rho$ where ρ is the fraction of the sample used for parameter estimation and Z_1 and Z_2 are independent standard normal random variables if a quadratic loss function is used.

When testing for equal forecast accuracy of nested models in finite sample instead, it is essential to keep into account that estimation error can cause the large model to produce less precise forecasts than the smaller with few parameters to estimate. In this light, Giacomini & White (2006) retain the effect of estimation errors and do not make assumptions on the positive-definiteness of the long run variance allowing for comparison of both nested and non nested model in a finite sample setting. Their tests take as given two set of forecasts from two methods and assume that the parameters of the models are estimated using a rolling window of fixed length but their results are also valid for full sample forecasts. The distribution of such tests is approximated under the assumption that the observed sequence of forecasts gets large.

Consider two models that at time t are used to generate one step ahead forecasts $\hat{y}_{1t+1|t}$ and $\hat{y}_{2t+1|t}$ using a fixed windows of ω observation. Each forecast $\hat{y}_{it+1|t}$ is a function of the data $(z_t, z_{t-1}, \ldots, z_{t-\omega+1})$ and parameter estimates $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ and is denoted by $\hat{y}_{it+1|t}(\hat{\beta}_i)$. From these forecasts, losses are $L(\hat{y}_{it+1|t}(\hat{\beta}_i), y_{t+1})$ for $t = T_R, \ldots, T-1$. These can be used to compare the models' finite sample forecast accuracy evaluated at the current parameters $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$, using the null hypothesis:

$$H_0 : E\left[L(\hat{y}_{1t+1|t}(\hat{\beta}_{1t}), y_{t+1}) - L(\hat{y}_{2t+1|t}(\hat{\beta}_{2t}), y_{t+1})\right] = 0 \tag{3.15}$$

in which the effect of estimation errors does not vanish.

Giacomini & White (2006) establish, under a finite estimation window, that

$$UGW = \sqrt{\frac{1}{T_P \sigma^2}} \sum_{t=T_R}^{T-1} \left[d_{t+1}(\hat{\beta}_{1t}, \hat{\beta}_{2t}, y_{t+1}) - E[d]\right] \rightarrow_d N(0,1) \tag{3.16}$$

where $d_{t+1}(\hat{\beta}_{1t}, \hat{\beta}_{2t}, y_{t+1}) = L(f_{1t+1|t}(\hat{\beta}_{1t}), y_{t+1}) - L(f_{2t+1|t}(\hat{\beta}_{2t}), y_{t+1})$ measures the differential loss while the asymptotic variance σ^2 is given by

$$\sigma^2 = \lim_{T_P \to \infty} Var\left(\sqrt{\frac{1}{T_P}} \sum_{t=T_R}^{T-1} \left[d_{t+1}(\hat{\beta}_{1t}, \hat{\beta}_{2t}, y_{t+1}) - E[d]\right]\right) \tag{3.17}$$

where T_R is the estimation sample size and T_P is the evaluation sample size.

This result is obtained in the limit for $T_p \to \infty$ but without the assumption that T_R expands asymptotically so there is no need for the variance to be positive definite. The UGW test statistic is equivalent to a Diebold and Mariano test statistic and fixed-smoothing asymptotics can also be used in this environment.

3.4 Estimates of the Long Run Variance

In order to test the null hypothesis of equal forecast accuracy we need our Diebold and Mariano test statistics, or other equivalent tests statistics, to be feasible so we need an estimate of the long run variance of the sample mean of the loss differential σ^2.

If a consistent estimate is used in the test statistic, the limit normality of the test statistic is not affected.

However, some methods may generate negative estimates, especially in small samples, and this can, in turn, cause serious size distortion.

In this chapter a series of estimation methods are presented: after the standard Weighted Covariance Estimator used in Diebold & Mariano (1995), other Heteroscedasticity-autocorrelation robust, or HAR, approaches such as fixed b and fixed m are discussed.

3.4.1 Consistent Weighted Autocovariance Estimation

In this case, the consistent estimator for the variance is obtained as a weighted sum of sample autocovariances (WCE) using a rectangular lag window using only $h-1$ sample autocovariances in the estimation of the spectral density of the loss differential at frequency 0, because all the others covariances are zero if the forecasts are optimal.

The estimator is:

$$\hat{\sigma}^2 = 2\pi \hat{f}_d(0) = \sum_{\tau=-M(T)}^{M(T)} k\left(\frac{\tau}{M(T)}\right)\hat{\gamma}_\tau = \hat{\gamma}_0 + 2\sum_{\tau=1}^{T-1} k\left(\frac{\tau}{M(T)}\right)\hat{\gamma}_\tau \tag{3.18}$$

where $\hat{\gamma}_\tau = \frac{1}{T}\sum_{t=|\tau|}^{T}(d_t - \bar{d})(d_{t-|\tau|} - \bar{d})$, $1 \leq M(T) \leq T$ is the bandwidth or lag truncation which increases with T but at a slower rate and $k(.)$ is the weighting scheme or kernel function.

For consistency of the variance estimator, regularity conditions include $M \to \infty$ and $\frac{M}{T} \to 0$ as $T \to \infty$.

In Diebold & Mariano (1995) the weighting scheme is the truncated uniform/rectangular kernel; the indicator function takes the value of unity when the argument has an absolute value less than one:

$$k\left(\frac{\tau}{M(T)}\right) = I\left(\left|\frac{\tau}{M(T)}\right| < 1\right)$$

To support this choice, they argue that in practice the h-step-ahead forecast errors are at most $h-1$ dependent even though this can be violated in some cases; setting $M = h-1$ means that only $h-1$ sample autocovariances will be used in the estimation of $f_d(0)$ as they assume all the others are zero.

The lag window is therefore defined as

$$I\left(\frac{\tau}{M(T)}\right) = 1 \quad for \quad \left|\frac{\tau}{M(T)}\right| \leq 1$$

and zero otherwise; as a result, the variance estimator of Diebold and Mariano can be written as

$$\hat{\sigma}^2_{DM} = \hat{\gamma}_0 + 2\sum_{\tau=1}^{h-1}\hat{\gamma}_\tau \tag{3.19}$$

This estimator is consistent as long as $\hat{\gamma}_0$ and $\hat{\gamma}_\tau$ are consistent ensuring:

$$\sqrt{T}\frac{\bar{d}-\mu}{\hat{\sigma}_{DM}} \quad \xrightarrow{d} \quad N(0,1) \quad under \quad H_0 \tag{3.20}$$

Unfortunately, the spectral density function estimator may not be positive semidefinite as the Dirichlet spectral window associated with the rectangular lag window plunges below zero at certain locations. In the unlikely event of a non positive semidefinite estimate, Diebold and Mariano suggest to treat the estimate as zero and automatically reject the null hypothesis, they also suggest using other spectral windows but simulations in Clark (1999) using a Bartlett kernel do not improve results.

3.4.2 Fixed-smoothing asymptotics

For consistency of the variance estimator in (3.19), regularity conditions are $M \to \infty$ and $\frac{M}{T} \to 0$ as $T \to \infty$; in order to satisfy these conditions, assuming the truncation lag M increases as T but at a slower rate is needed. However, in practice, the sample size is given and a positive fraction is used for a particular data set.
Neave (1970) argues that:

> This is a convenient assumption mathematically in that, in particular, it ensures consistency of the estimates, but it is unrealistic when such results are used as approximations to the finite case where the value of $\frac{M}{T}$ cannot be zero.

Then, he showed that assuming that $\frac{M}{T}$ is a constant, it gives more accurate results than the standard approximation.
Kiefer & Vogelsang (2005) generalise the approach of Neave (1970) for zero frequency non parametric spectral density estimators, they consider $b = \frac{M}{T}$ as fixed and $b \in (0; 1]$, this approach is referred to as fixed b asymptotics in contrast to the standard asymptotics where b goes to zero as T increases also known as small b asymptotics.

In standard asymptotics the effect of kernel function and bandwidth M are not captured in the limiting distribution while, in fixed b asymptotics, the limiting distribution depends on the kernel and b.

Fixed b approach is a better reflection of practice in reality and leads to a more accurate first-order approximation for tests: the error in rejection probability is of a smaller order than the one of the standard normal approximation.

Assuming that the Bartlett kernel (Bartlett, 1950) is used, the estimator for the variance is

$$\hat{\sigma}^2_{BART} = \hat{\gamma}_0 + 2\sum_{\tau=1}^{T-1} k_{BART}\left(\frac{\tau}{M(T)}\right)\hat{\gamma}_\tau \tag{3.21}$$

where

$$k_{BART}\left(\frac{\tau}{M(T)}\right) = \begin{cases} 1 - \left|\frac{\tau}{M(T)}\right| & \text{if } \left|\frac{\tau}{M(T)}\right| \leq 1; \\ 0 & \text{otherwise} \end{cases} \tag{3.22}$$

The asymptotic distribution of the test is

$$\sqrt{T}\frac{\bar{d} - \mu}{\hat{\sigma}_{BART}} \Longrightarrow \Phi_{BART}(b) \tag{3.23}$$

where $\Longrightarrow$ denotes weak convergence in the $D[0,1]$ space with Skorohod topology and $\Phi_{BART}(b)$ is characterised in Kiefer & Vogelsang (2005).

As the limiting distribution is non standard, they provide critical value using the cubic equation

$$cv(b) = a_0 + a_1 b + a_2 b^2 + a_3 b^3 \tag{3.24}$$

and a table for coefficients according to kernel choice and percentage points of 90%, 95%, 97.5% and 99% via simulation methods.

Finally, their simulations show there is a trade-off between size distortion and power: large b leads to a small size distortion but at the cost of low power and vice-versa.

Similar trade-offs appears across different kernel choices.

Fixed b asymptotics provide a more accurate approximation for the variance estimator with a limiting distribution which depends on both kernel and b. The limiting distribution is non standard but given that it is a function of Brownian motions, critical values can be easily obtained using simulations, in addition Kiefer & Vogelsang (2005) provide the formula (3.24) to calculate critical values.

Sun (2013), Müller (2014) and Hualde & Iacone (2015) among others, suggest a long run variance estimator based on Weighted Periodogram Estimation closely related to the one obtained in the fixed b case but that does not require simulations because, for weakly dependent processes, the standardized mean has a t student distribution as shown in similar cases by Sun (2013, Theorem 3.1) and Müller (2014, p. 314).

Recall that the loss differential for two forecast defined in (3.1) can be expressed as $d_t = \bar{d} + \hat{u}_t$ with unknown long run variance $\sigma^2 = \sum_{\tau=-\infty}^{+\infty} \gamma_\tau$ where $\gamma_\tau = E[u_t u_{t+\tau}]$ and $E[u_t] = 0$.

Moving from the time domain to the frequency domain, now consider the Fourier transform of d_t with $\lambda_\tau := \frac{2\pi\tau}{T}$ as the Fourier frequencies

$$w_d(\lambda_\tau) := \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} d_t e^{i\lambda_\tau t} \tag{3.25}$$

and its periodogram

$$I(\lambda_\tau) := |w_d(\lambda_\tau)|^2 \tag{3.26}$$

which is also the periodogram of $\hat{u}_t$ at these frequencies as $\frac{1}{\sqrt{2\pi T}} \bar{d} \sum_{t=1}^{T} e^{i\lambda_\tau t} = 0$, given that $\sum_{t=1}^{T} e^{i\lambda_\tau t} = 0$ for $\tau \neq 0$.

Weighted periodogram estimation methods provide an estimator for the long run variance of the form

$$\tilde{\sigma}^2 = 2\pi \sum_{\tau=1}^{T/2} K_M(\lambda_\tau) I(\lambda_\tau) \tag{3.27}$$

where K_M is the spectral window.

For $K_M := (2\pi)^{-1} \sum_{|\tau|<T} k(\tau/M) e^{-i\tau\lambda}$, the Weighted Covariance Estimate $\sum_{\tau=-M}^{M} k\left(\frac{\tau}{M}\right)\hat{\gamma}_\tau$ has frequency domain representation

$$\int_{-\pi}^{\pi} K_M(\lambda_\tau) I^*(\lambda_\tau) d\lambda \tag{3.28}$$

where $I^*(\lambda_\tau)$ is the periodogram of $d_t - \bar{d}$.

Using the Daniell kernel (Daniell, 1946)

$$k(x) := \frac{\sin(x)}{x} \quad \forall x \tag{3.29}$$

the spectral window $K_M(\lambda_\tau)$ takes value

$$K_M(\lambda_\tau) = \begin{cases} \frac{M}{2\pi} & \text{if } -\frac{\pi}{M} \leq \lambda_\tau \leq \frac{\pi}{M}; \\ 0 & \text{otherwise} \end{cases} \tag{3.30}$$

Given the form of the spectral window, Weighted Covariance Estimation and Weighted Periodogram Estimation are related, in fact, the Weighted Periodogram Estimate of σ^2

$$\hat{\sigma}^2_{DAN} = 2\pi \frac{1}{m} \sum_{\tau=1}^{m} I(\lambda_\tau), \tag{3.31}$$

for $m = T/(2M)$, a function of the bandwidth, is an approximation of (3.28) when the Daniell kernel is used. If $m \to \infty$, $\hat{\sigma}^2_{DAN}$ is consistent; if m is fixed, it is not consistent but it is still asymptotically unbiased.

The test statistic for the null hypothesis of equal predictive accuracy, considering m

fixed, has limit distribution

$$\sqrt{T}\frac{\bar{d}-\mu}{\hat{\sigma}_{DAN}} \xrightarrow{d} t_{2m} \tag{3.32}$$

As fixed b and fixed m approaches are very similar, they both suffer from size-power trade off: the smaller the value for m (the larger b, at least for WCE) the better the empirical size, but also the weaker power.

3.5 Montecarlo Study

In this section, I perform a Montecarlo study to assess size and power performances of the Diebold and Mariano test and the Unconditional Giacomini and White test paired with fixed smoothing asymptotics.

3.5.1 Diebold and Mariano Test

Starting from a vector of forecast error innovations from a bivariate standard normal $(v_{1t}, v_{2t})' \sim N(0_2, I_2)$, I introduce contemporaneous correlation taking

$$\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \begin{pmatrix} \sqrt{k} & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix} \tag{3.33}$$

with $k = 1$ and $\rho = 0.5$.

Forecast errors are heteroscedastic and constructed as ARCH(1)

$$e_{1t} = \sigma_{1t} u_{1t} \tag{3.34}$$

where $\sigma_{1t}^2 = \alpha_0 + \alpha_1 e_{1t-1}^2$, $\alpha_0 = 0.7$, $\alpha_1 = 0.1$,

$$e_{2t} = \sigma_{2t} u_{2t} \tag{3.35}$$

where $\sigma_{2t}^2 = \beta_0 + \beta_1 e_{2t-1}^2$ and $\beta_0 = \alpha_0$, $\beta_1 = \alpha_1$ for H_0 to be true.

The theoretical size is set to 5% and I use $10,000$ replications with a quadratic loss function.

WCE estimates are performed using the Diebold and Mariano estimator in (3.19) and the one in (3.21) with $M = T^{1/3}$, $M = T^{1/2}$ and $M = T$. WPE estimates follow (3.31) with $m = T^{1/4}$, $m = T^{1/3}$, $m = T^{1/2}$ and $m = T^{2/3}$.

Results are reported in Table 3.1. As the sample size increases, the size distortion reduces and, in all cases, there is a strong improvement from using fixed-smoothing asymptotics instead of standard asymptotics.

For the standard Diebold & Mariano Test, size is better the larger the bandwidth for WCE and the smaller the bandwidth for WPE confirming the findings of Coroneo & Iacone (2015). The best bandwidths for both WCE and WPE are $M = T^{1/3}$ and $m = T^{2/3}$ respectively.

For this power analysis, I only consider estimates of the long run variance made using Weighted Covariance Estimators with Bartlett kernel and Weighted Periodogram Estimators with Daniell kernel and critical values taken from fixed-smoothing asymptotics. Forecast innovations are generated by a bivariate Normal distribution $(u_{1t}, e_{2t})' \sim N(0_2, I_2)$ and forecast errors are constructed as in the size exercise. The theoretical size is set to 5% and the exercise involves $10,000$ replications.

The loss differential is quadratic and $cT^{-1/2}$ is added to generate local alternatives and obtain $\mu = cT^{-1/2}$ testing the null hypothesis $H_0 : \mu = 0$. Results are displayed in Figure 3.1, the 'U' case is the unfeasible case obtained using the true variance.

Also in this case, results are quite encouraging, although a size/power trade off is

Table 3.1: Empirical Size of DM Test with Standard and Fixed-smoothing asymptotics. Powers of T are bandwidths for WCE and WPE and theoretical size is 5%. The long run variance for the column 'DM' is estimated as in the original paper while in other WCE cases using a Bartlett kernel and using a Daniell kernel in WPE cases.

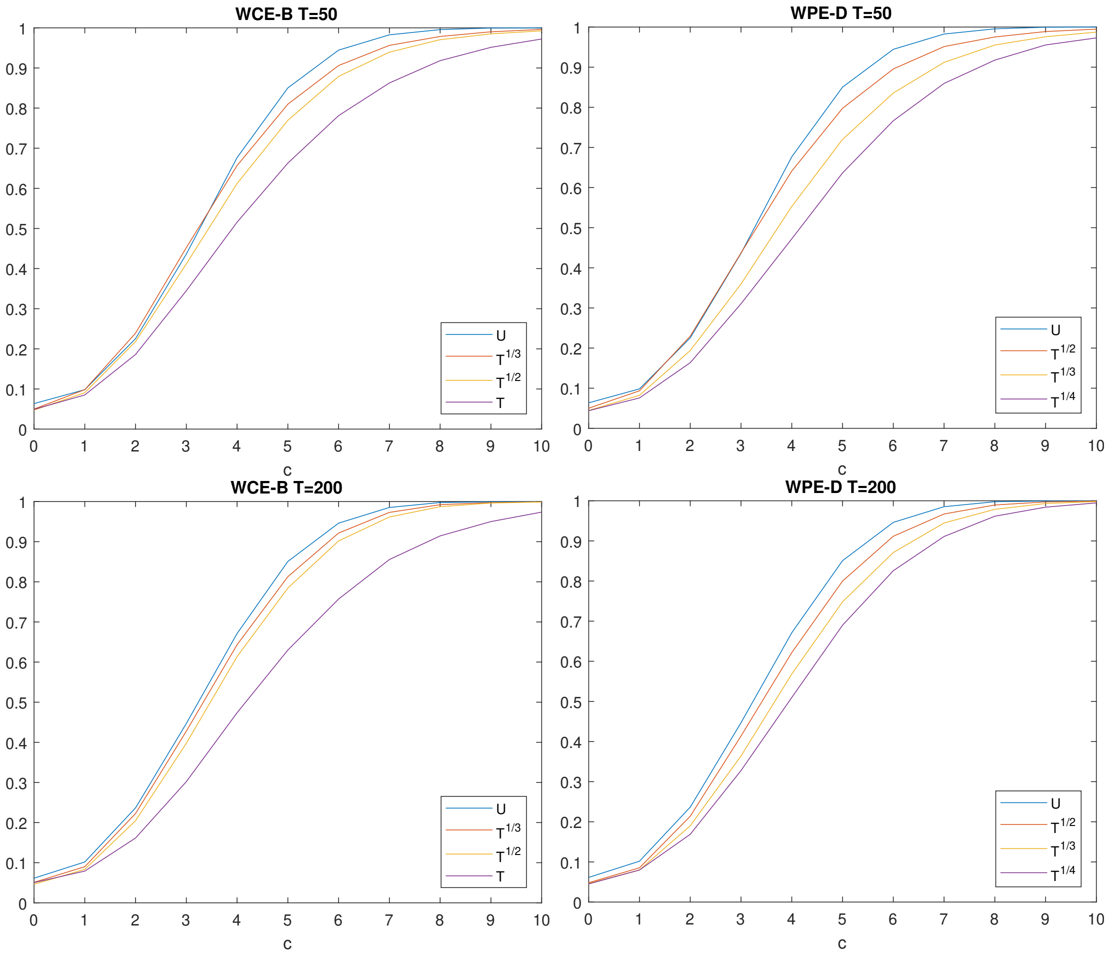
T=50

Standard Asymptotics							
WCE				**WPE**			
DM	$T^{1/3}$	$T^{1/2}$	T	$T^{1/4}$	$T^{1/3}$	$T^{1/2}$	$T^{2/3}$
0.070	0.075	0.099	0.349	0.107	0.089	0.065	0.066
Fixed-smoothing asymptotics							
WCE				**WPE**			
	$T^{1/3}$	$T^{1/2}$	T	$T^{1/4}$	$T^{1/3}$	$T^{1/2}$	$T^{2/3}$
	0.047	0.043	0.047	0.046	0.041	0.045	0.052

T=200

Standard Asymptotics							
WCE				**WPE**			
DM	$T^{1/3}$	$T^{1/2}$	T	$T^{1/4}$	$T^{1/3}$	$T^{1/2}$	$T^{2/3}$
0.053	0.059	0.075	0.336	0.097	0.076	0.059	0.054
Fixed-smoothing asymptotics							
WCE				**WPE**			
	$T^{1/3}$	$T^{1/2}$	T	$T^{1/4}$	$T^{1/3}$	$T^{1/2}$	$T^{2/3}$
	0.051	0.048	0.047	0.047	0.049	0.049	0.050

Figure 3.1: Empirical rejection frequencies at 5% nominal size. The 'U' case is the unfeasible case obtained using the unknown true variance. Estimates of the long run variance are obtained from WCE with Bartlett kernel and from WPE with Daniell kernel.



present, confirming findings from Coroneo & Iacone (2015) and Harvey et al. (2016) in case of serial correlation only. The best size and power performances are obtained with bandwidths $M = T^{1/3}$ and $m = T^{1/2}$ for WCE and WPE respectively which seems to provide correctly sized test statistics without loosing too much power.

3.5.2 Giacomini and White Unconditional Test

In this section, I investigate the size and power behaviour of the Giacomini and White unconditional test when paired with fixed-smoothing asymptotics. The setting of this Montecarlo is similar to the one presented in Giacomini & White (2006). The data generating process is obtained from the second log difference of the monthly US consumer price index from 1959.1 to 1998.12 (CPI_t):

$$Y_t = c + CPI_t + \epsilon, \quad \epsilon \sim iid\ N(0, \sigma^2) \tag{3.36}$$

The two competing models are

$$Y_t = \beta CPI_t + u_{1t} \tag{3.37}$$

as the misspecified model in which the intercept is omitted, and

$$Y_t = \delta + \gamma CPI_t + u_{2t} \tag{3.38}$$

as the correctly specified model.

One step ahead forecasts are obtained as

$$\hat{y}^{(1)}_{t+1} = \hat{\beta}_t CPI_{t+1} \tag{3.39}$$

and

$$\hat{y}^{(2)}_{t+1} = \hat{\delta}_t + \hat{\gamma}_t CPI_{t+1}. \tag{3.40}$$

Using the result reported in Giacomini & White (2006, Proposition 5), for the two forecasting models to have equal expected Mean Squared Error, c must satisfy the following equation:

$$c = \sigma \left(\left\{ \sum_{t=m}^{T-1} \left(\left(\sum_{j=t-m+1}^{t} CPI_j^2 / mS \right) + CPI_{t+1}^2 / S - 2(\overline{CPI}/S) CPI_{t+1} / \sum_{j=t-m+1}^{t} CPI_j^2 \right) \right\} \times \left\{ \sum_{t=m}^{T-1} \left(1 - \left(\sum_{j=t-m+1}^{t} CPI_j / \sum_{j=t-m+1}^{t} CPI_t^2 \right) CPI_{t+1} \right)^2 \right\}^{-1} \right)^{1/2} \tag{3.41}$$

where $\overline{CPI} = 1/m \sum_{j=t-m+1}^{t} CPI_j$, $S = \sum_{j=t-m+1}^{t} CPI_j^2 - m\overline{CPI}^2$, $\sigma = 0.1$, T is the total sample size and m is the rolling window estimation sample size fixed at 100 in this exercise.

As in the Diebold and Mariano case, the theoretical size is 5% and I use 10,000 replications with a quadratic loss function. WCE estimates are performed using the Diebold and Mariano estimator in (3.19) and the one in (3.21) with bandwidths $M = T^{1/3}$, $M = T^{1/2}$ and $M = T$. WPE estimates follow (3.31) with bandwidths $m = T^{1/4}$, $m = T^{1/3}$, $m = T^{1/2}$ and $m = T^{2/3}$.

Results are reported in Table 3.2. The original test appears to be slightly undersized, standard asymptotics provide oversized test statistics for some bandwidths such as $M = T$ and $m = T^{1/4}$ for WCE and WPE respectively, while fixed-smoothing asymptotics provide quite good results especially for bandwidth $M = T^{1/2}$ in the WCE case and $m = T^{1/3}$ in the WPE case.

For the power analysis, I directly generate a sequence of 50 and 300 loss differentials

Table 3.2: Empirical Size of UGW Test with Standard and Fixed-smoothing asymptotics. Powers of T are bandwidths and theoretical size is 5%. The long run variance for the column 'UGW' is estimated as in the original paper while in other WCE cases using a Bartlett kernel and using a Daniell kernel in WPE cases.

T=376

Standard Asymptotics							
WCE				**WPE**			
UGW	$T^{1/3}$	$T^{1/2}$	T	$T^{1/4}$	$T^{1/3}$	$T^{1/2}$	$T^{2/3}$
0.037	0.042	0.058	0.334	0.093	0.070	0.046	0.038
Fixed-smoothing Asymptotics							
WCE				**WPE**			
	$T^{1/3}$	$T^{1/2}$	T	$T^{1/4}$	$T^{1/3}$	$T^{1/2}$	$T^{2/3}$
	0.037	0.043	0.057	0.060	0.055	0.040	0.036

following

$$d_{t+1} = \rho d_t + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim iid \;\; N(0,1). \tag{3.42}$$

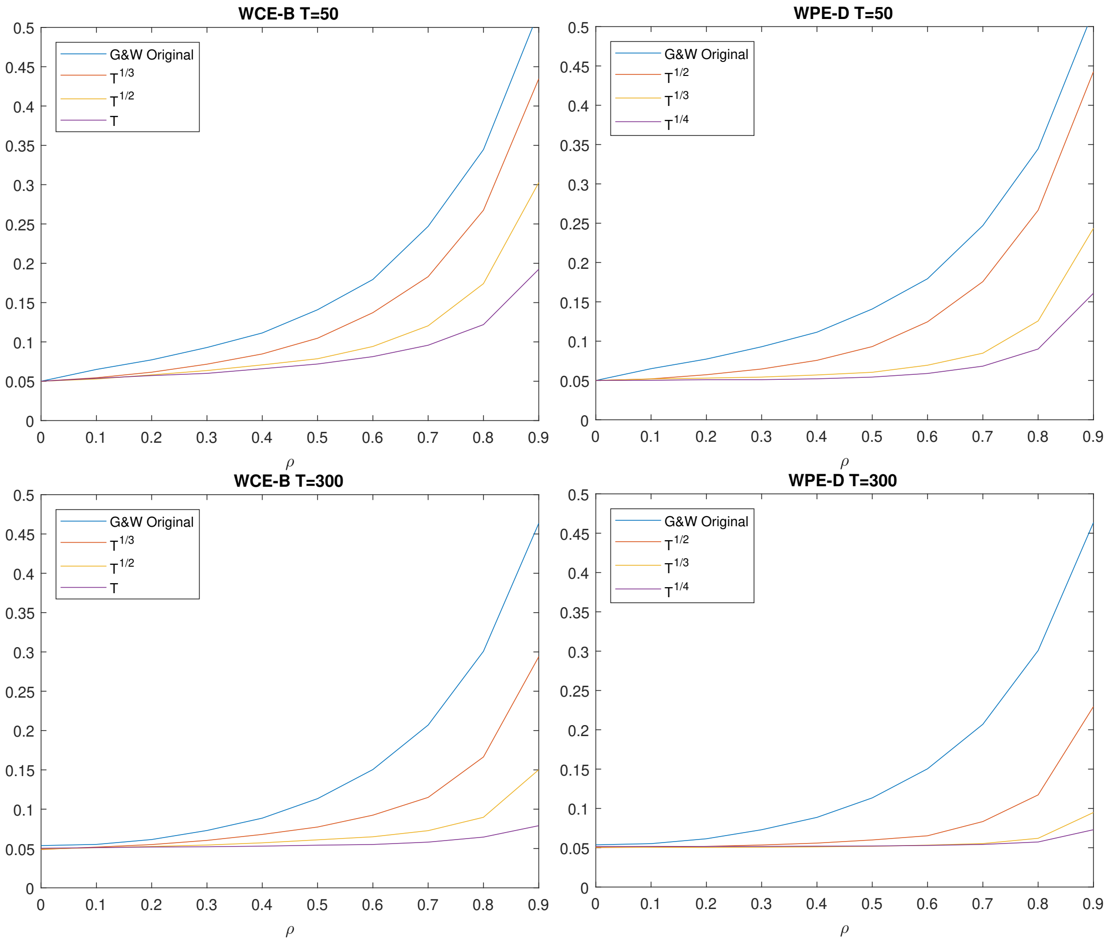
for each 10,000 Montecarlo replications. ρ indicates how the loss is serially correlated and spans from 0 to 0.9.

Results are reported in Figure 3.2 for sample sizes of 50 and 300. Confirming results from Giacomini & White (2006), the unconditional test suffers from severe size distortion as serial correlation increases but fixed smoothing asymptotics improve results and reduce distortion as can be seen from the fact that the curve obtained with alternative asymptotics tends to remain flat and close to the theoretical 0.05 size.

3.6 Empirical Example

Along the lines of the original Diebold & Mariano (1995) paper, my empirical application is about exchange rate forecasting. I consider forecasts of the monthly

Figure 3.2: Empirical rejection frequencies at 5% nominal size. The 'G&W Original' case is the one obtained using the original test specifications. Estimates of the long run variance are obtained from WCE with Bartlett kernel and from WPE with Daniell kernel.



three-month change in the nominal Dollar/Euro end-of-month spot exchange rate from July 1999 to June 2018. The first competing forecast is from a random walk model which is constant at zero. The second forecast I consider is given by the difference between the three-month forward rate and the spot rate.

Considering that the Diebold and Mariano test is simple, easy to compute and still widely used even when forecasts are obtained from estimated models (Diebold, 2015), my empirical application is based on it. Also, in this setting in which forecasts are deemed model-free, the Diebold and Mariano test seems the most appropriate to use among those available.

Table 3.3: DM Test for equal forecast accuracy. Critical values are for two sided tests.

Test Type	**Test Statistic**	**10% CV**	**5% CV**
DM	-6.522	1.645	1.960
WCE-B, $M = T^{1/3}$	-6.866	1.703	2.038
WCE-B, $M = T^{1/2}$	-6.041	1.790	2.157
WPE-D, $m = T^{1/4}$	-4.631	1.943	2.447
WPE-D, $m = T^{1/3}$	-5.381	1.782	2.179

Table 3.3 shows results for the Diebold and Mariano test of equal forecast accuracy with a quadratic loss function. The null hypotesis is strongly rejected in favour of the random walk forecast with all kind of long run variance estimators described in this chapter and fixed smoothing asymptotics critical values confirming the findings by Diebold & Mariano (1995).

3.7 Conclusions

In this chapter, I use fixed-smoothing asymptotics to overcome size distortion in Diebold & Mariano (1995) and equivalent tests, such as the Giacomini & White (2006) one, in case of heteroscedasticity and serial correlation. My Montecarlo exercise shows these alternative asymptotics provide correctly sized test statistics even when sample size is quite small. Confirming results from the literature, size performances are improving but a size/power trade off arises.

As empirical example of the application of the Diebold and Mariano test, I decided to recreate an application very similar to the one in the original Diebold & Mariano

(1995) paper. My findings confirm what it is well established in the literature: random walk forecasts are hard to beat.

As these results seem promising, fixed smoothing asymptotics could be applied in other tests in which a long run variance estimator is needed such as White (2000)'s Reality Check and Hansen (2005)'s Superior Predictive Ability test.

Bibliography

Ammann, M., Skovmand, D., & Verhofen, M. (2009). Implied and realized volatility in the cross-section of equity options. *International Journal of Theoretical and Applied Finance*, *12*(06), 745–765.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2002). Parametric and nonparametric volatility measurement.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical finance*, *9*(3), 203–228.

Asplund, M., Eriksson, R., & Friberg, R. (2000). Price adjustments by a gasoline retail chain. *Scandinavian Journal of Economics*, *102*(1), 101–121.

Baba, Y., Engle, R., Kraft, D., & Kroner, K. (1990). Multivariate simultaneous generalized ARCH, department of economics, university of california at san diego. Technical report, Working Paper.

Bachmeier, L. J. & Griffin, J. M. (2003). New evidence on asymmetric gasoline price responses. *The Review of Economics and Statistics*, *85*(3), 772–776.

Bacon, R. W. (1991). Rockets and feathers: the asymmetric speed of adjustment of uk retail gasoline prices to cost changes. *Energy economics*, *13*(3), 211–218.

Bailey, D. & Brorsen, B. W. (1989). Price asymmetry in spatial fed cattle markets. *Western Journal of Agricultural Economics*, *14*(2), 246–252.

Baillie, R. T. & DeGennaro, R. P. (1990). Stock returns and volatility. *Journal of financial and Quantitative Analysis*, *25*(2), 203–214.

Baillie, R. T. & Myers, R. J. (1991). Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics*, *6*(2), 109–124.

Balke, N. S., Brown, S. P., & Yucel, M. K. (1998). Crude oil and gasoline prices: an asymmetric relationship? *Economic Review-Federal Reserve Bank of Dallas*, *issue Q*(1), 2–11.

Balmaceda, F. & Soruco, P. (2008). Asymmetric dynamic pricing in a local gasoline retail market. *The Journal of Industrial Economics*, *56*(3), 629–653.

Bartlett, M. S. (1950). Periodogram analysis and continuous spectra. *Biometrika*, *37*(1/2), 1–16.

Batten, J. A. & In, F. (2006). Dynamic interaction and valuation of quality yen eurobonds in a multivariate EGARCH framework. *Applied Financial Economics*, *16*(12), 881–892.

Bkedowska-Sojka, B. & Kliber, A. (2010). Realized volatility versus GARCH and stochastic volatility models. the evidence from the wig20 index and the eur/pln foreign exchange market. *Przeglkad Statystyczny*, *57*(4), 105–127.

Blinder, A., Canetti, E. R., Lebow, D. E., & Rudd, J. B. (1998). *Asking about prices: a new approach to understanding price stickiness*. Russell Sage Foundation.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *The review of economics and statistics*, *72*(3), 498–505.

Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *Journal of political Economy*, *96*(1), 116–131.

Bonnet, C. & Villas-Boas, S. B. (2016). An analysis of asymmetric consumer price responses and asymmetric cost pass-through in the french coffee market. *European Review of Agricultural Economics*, *43*(5), 781–804.

Borenstein, S., Cameron, A. C., & Gilbert, R. (1997). Do gasoline prices respond asymmetrically to crude oil price changes? *The Quarterly Journal of Economics*, *112*(1), 305–339.

Borenstein, S. & Shepard, A. (1996). Sticky prices, inventories, and market power in wholesale gasoline markets. Technical report, National Bureau of Economic Research.

Borenstein, S. & Shepard, A. (2002). Sticky prices, inventories and market power in wholesale gasoline markets. *RAND Journal of Economics*, *33*(1), 116–139.

Brooks, C., Henry, O. T., & Persand, G. (2002). The effect of asymmetries on optimal hedge ratios. *The Journal of Business*, *75*(2), 333–352.

Cecchetti, S. G., Cumby, R. E., & Figlewski, S. (1988). Estimation of the optimal futures hedge. *The Review of Economics and Statistics*, *70*(4), 623–630.

Chesnes, M. (2016). Asymmetric pass-through in us gasoline prices. *Energy Journal*, *37*(1), 153–180.

Christoffersen, P. F. & Diebold, F. X. (1996). Further results on forecasting and model selection under asymmetric loss. *Journal of applied econometrics*, *11*(5), 561–571.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, *41*(1), 135–155.

Clark, T. E. (1999). Finite-sample properties of tests for equal forecast accuracy. *Journal of Forecasting*, *18*(7), 489–504.

Clark, T. E. & McCracken, M. W. (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of econometrics*, *105*(1), 85–110.

Clark, T. E. & McCracken, M. W. (2005). Evaluating direct multistep forecasts. *Econometric Reviews*, *24*(4), 369–404.

Clements, M. P. & Hendry, D. F. (1993). On the limitations of comparing mean square forecast errors. *Journal of Forecasting*, *12*(8), 617–637.

Copeland, L. & Zhu, Y. (2006). Hedging effectiveness in the index futures market. Cardiff Economics Working Papers E2006/10, Cardiff University, Cardiff Business School, Economics Section.

Coroneo, L. & Iacone, F. (2015). Comparing predictive accuracy in small samples. Technical Report 15/15, The University of York.

Cramon-Taubadel, S. V. (1998). Estimating asymmetric price transmission with the error correction representation: An application to the german pork market. *European review of agricultural economics*, *25*(1), 1–18.

Daniell, P. J. (1946). Discussion on symposium on autocorrelation in time series. *Supplement to Journal of the Royal Statistical Society*, *8*, 88–90.

Deltas, G. (2008). Retail gasoline price dynamics and local market power. *The Journal of Industrial Economics*, *56*(3), 613–628.

Diebold, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of diebold–mariano tests. *Journal of Business & Economic Statistics*, *33*(1), 1–1.

Diebold, F. X. & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & economic statistics*, *13*(3), 253–262.

Duffy-Deno, K. T. (1996). Retail price asymmetries in local gasoline markets. *Energy Economics*, *18*(1-2), 81–92.

Durbin, J. & Watson, G. S. (1950). Testing for serial correlation in least squares regression: I. *Biometrika*, *37*(3/4), 409–428.

Durbin, J. & Watson, G. S. (1951). Testing for serial correlation in least squares regression. ii. *Biometrika*, *38*(1/2), 159–177.

Eckert, A. (2002). Retail price cycles and response asymmetry. *Canadian Journal of Economics/Revue canadienne d'économique*, *35*(1), 52–77.

Eckert, A. (2013). Empirical studies of gasoline retailing: A guide to the literature. *Journal of Economic Surveys*, *27*(1), 140–166.

Ederington, L. H. (1979). The hedging performance of the new futures markets. *The Journal of Finance*, *34*(1), 157–170.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, *20*(3), 339–350.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, *50*(4), 987–1007.

Engle, R. F. & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, *55*(2), 251–276.

Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica: journal of the Econometric Society*, *55*(2), 391–407.

Farrell, M. J. (1952). Irreversible demand functions. *Econometrica*, *20*(2), 171–186.

Freebairn, J. W. (1984). Farm and retail food prices. *Review of Marketing and Agricultural Economics*, *52*(2), 71–90.

Frey, G. & Manera, M. (2007). Econometric models of asymmetric price transmission. *Journal of Economic surveys*, *21*(2), 349–415.

Galeotti, M., Lanza, A., & Manera, M. (2003). Rockets and feathers revisited: an international comparison on european gasoline markets. *Energy economics*, *25*(2), 175–190.

Ghosh, A. (1993a). Cointegration and error correction models: Intertemporal causality between index and futures prices. *Journal of futures markets*, *13*(2), 193–198.

Ghosh, A. (1993b). Hedging with stock index futures: Estimation and forecasting with error correction model. *Journal of Futures Markets*, *13*(7), 743–752.

Ghosh, A. & Clayton, R. (1996). Hedging with international stock index futures: an intertemporal error correction model. *Journal of Financial Research*, *19*(4), 477–491.

Giacomini, R. & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, *74*(6), 1545–1578.

Girapunthong, N., VanSickle, J. J., & Renwick, A. (2003). Price assymetry in the united states fresh tomato market. *Journal of Food Distribution Research*, *34*(3), 51–59.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, *48*(5), 1779–1801.

Granger, C. W. & Pesaran, M. H. (2004). Economic and statistical measures of forecast accuracy. Technical report, University of Cambridge.

Granger, C. W. J. & Newbold, P. (1986). *Forecasting economic time series*. Academic Press.

Grasso, M. & Manera, M. (2007). Asymmetric error correction models for the oil–gasoline price relationship. *Energy Policy*, *35*(1), 156–177.

Greenwood-Nimmo, M. & Shin, Y. (2013). Taxation and the asymmetric adjustment of selected retail energy prices in the uk. *Economics Letters*, *121*(3), 411–416.

Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics*, *23*(4), 365–380.

Hansen, P. R. & Timmermann, A. (2012). Choice of sample split in out-of-sample forecast evaluation. Technical report.

Harvey, D. (1997). *The evaluation of economic forecasts.* PhD thesis, University of Nottingham.

Harvey, D. I., Leybourne, S. J., & Whitehouse, E. J. (2016). Testing forecast accuracy in small samples. Technical report, University of Nottingham.

Hatemi-J, A. & Roca, E. (2006). Calculating the optimal hedge ratio: constant, time varying and the kalman filter approach. *Applied Economics Letters*, *13*(5), 293–299.

Heien, D. M. (1980). Markup pricing in a dynamic model of the food industry. *American Journal of Agricultural Economics*, *62*(1), 10–18.

Hogg, R. V. & Craig, A. T. (1970). *Introduction to Mathematical Statistics: 3d Ed.* Macmillan.

Hualde, J. & Iacone, F. (2015). Autocorrelation robust inference using the daniell kernel with fixed bandwidth. Technical Report 15/14, The University of York.

Jarque, C. M. & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, *6*(3), 255–259.

Johnson, R. N. (2002). Search costs, lags and prices at the pump. *Review of Industrial Organization*, *20*(1), 33–50.

Jorion, P. (2000). *Value-at-risk: The New Benchmark for Managing Financial Risk.* McGraw-Hill, New York.

Karrenbrock, J. D. (1991). The behavior of retail gasoline prices: symmetric or not? *Federal Reserve Bank of St. Louis Review*, *73*(4), 19–29.

Kaufmann, R. K. & Laskowski, C. (2005). Causes for an asymmetric relation between the price of crude oil and refined petroleum products. *Energy Policy*, *33*(12), 1587–1596.

Kenourgios, D., Samitas, A., & Drosos, P. (2008). Hedge ratio estimation and hedging effectiveness: the case of the s&p 500 stock index futures contract. *International Journal of Risk assessment and management*, *9*(1-2), 121–134.

Kiefer, N. M. & Vogelsang, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory*, *21*(6), 1130–1164.

Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *The review of economic studies*, *65*(3), 361–393.

Kinnucan, H. W. & Forker, O. D. (1987). Asymmetry in farm-retail price transmission for major dairy products. *American journal of agricultural economics*, *69*(2), 285–292.

Koenker, R. (1981). A note on studentizing a test for heteroscedasticity. *Journal of Econometrics*, *17*(1), 107–112.

Kroner, K. F. & Sultan, J. (1990). *Exchange rate volatility and time varying hedge ratios*. Bentley College, Institute for Research and Faculty Development.

Kroner, K. F. & Sultan, J. (1993). Time-varying distributions and dynamic hedging with foreign currency futures. *Journal of financial and quantitative analysis*, *28*(4), 535–551.

Lehar, A., Scheicher, M., & Schittenkopf, C. (2002). GARCH vs. stochastic volatility: Option pricing and risk management. *Journal of Banking & Finance*, *26*(2-3), 323–345.

Lewis, M. S. (2011). Asymmetric price adjustment and consumer search: An examination of the retail gasoline market. *Journal of Economics & Management Strategy*, *20*(2), 409–449.

Lien, D. (2009). A note on the hedging effectiveness of GARCH models. *International Review of Economics & Finance*, *18*(1), 110–112.

Lien, D. (2010). A note on the relationship between the variability of the hedge ratio and hedging performance. *Journal of Futures Markets*, *30*(11), 1100–1104.

Lien, D.-H. D. (1996). The effect of the cointegration relationship on futures hedging: a note. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, *16*(7), 773–780.

Manning, D. (1991). Petrol prices, oil price rises and oil price falls: some evidence for the uk since 1972. *Applied economics*, *23*(9), 1535–1541.

McCracken, M. W. (2007). Asymptotics for out of sample tests of granger causality. *Journal of Econometrics*, *140*(2), 719–752.

Meese, R. & Rogoff, K. (1988). Was it real? the exchange rate-interest differential relation over the modern floating-rate period. *The Journal of Finance*, *43*(4), 933–948.

Meyer, J. & von Cramon-Taubadel, S. (2004). Asymmetric price transmission: a survey. *Journal of agricultural economics*, *55*(3), 581–611.

Miffre, J. (2004). Conditional ols minimum variance hedge ratios. *Journal of Futures Markets*, *24*(10), 945–964.

Morgan, W. (1939). A test for the significance of the difference between the two variances in a sample from a normal bivariate population. *Biometrika*, *31*(1/2), 13–19.

Müller, U. K. (2014). Hac corrections for strongly autocorrelated time series. *Journal of Business & Economic Statistics*, *32*(3), 311–322.

Murry, D. & Zhu, Z. (2008). Asymmetric price responses, market integration and market power: A study of the us natural gas market. *Energy Economics*, *30*(3), 748–765.

Neave, H. R. (1970). An improved formula for the asymptotic variance of spectrum estimates. *The Annals of Mathematical Statistics*, *41*(1), 70–77.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, *59*(2), 347–370.

Noel, M. (2009). Do retail gasoline prices respond asymmetrically to cost shocks? the influence of edgeworth cycles. *The RAND Journal of Economics*, *40*(3), 582–595.

Park, S. Y. & Jei, S. Y. (2010). Estimation and hedging effectiveness of time-varying hedge ratio: Flexible bivariate GARCH approaches. *Journal of Futures Markets*, *30*(1), 71–99.

Park, T. H. & Switzer, L. N. (1995). Bivariate GARCH estimation of the optimal hedge ratios for stock index futures: A note. *Journal of Futures Markets*, *15*(1), 61–67.

Peltzman, S. (2000). Prices rise faster than they fall. *Journal of political economy, 108*(3), 466–502.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, *16*(3), 289–326.

Pesaran, M. H. & Timmermann, A. (1994). Forecasting stock returns an examination of stock market trading in the presence of transaction costs. *Journal of Forecasting, 13*(4), 335–367.

Poterba, J. M., Rotemberg, J. J., & Summers, L. H. (1986). A tax-based test for nominal rigidities. *The American Economic Review*, *54*(6), 659–675.

Radchenko, S. (2005). Oil price volatility and the asymmetric response of gasoline prices to oil price increases and decreases. *Energy economics*, *27*(5), 708–730.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, *31*(2), 350–371.

Reilly, B. & Witt, R. (1998). Petrol price asymmetries revisited. *Energy Economics, 20*(3), 297–308.

Ruiz, E., Pena, D., & Carnero, M. A. (2001). Is stochastic volatility more flexible than GARCH? Technical report, Universidad Carlos III de Madrid. Departamento de Estadistica.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. *Monographs on Statistics and Applied Probability*, *65*, 1–68.

Shin, Y., Yu, B., & Greenwood-Nimmo, M. (2014). Modelling asymmetric cointegration and dynamic multipliers in a nonlinear ardl framework. In *Festschrift in Honor of Peter Schmidt* (pp. 281–314). Springer.

So, M. K., Lam, K., & Li, W. (1999). Forecasting exchange rate volatility using autoregressive random variance model. *Applied Financial Economics*, *9*(6), 583–591.

Stock, J. H. & Watson, M. W. (1999). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. Technical report, National Bureau of Economic Research.

Sun, Y. (2013). A heteroskedasticity and autocorrelation robust f test using an orthonormal series variance estimator. *The Econometrics Journal*, *16*(1), 1–26.

Tappata, M. (2009). Rockets and feathers: Understanding asymmetric pricing. *The RAND Journal of Economics*, *40*(4), 673–687.

Taylor, S. J. (1986). *Modelling financial time series.* John Wiley and Sons.

Verlinda, J. A. (2008). Do rockets rise faster and feathers fall slower in an atmosphere of local market power? evidence from the retail gasoline market. *The Journal of Industrial Economics*, *56*(3), 581–612.

Ward, R. W. (1982). Asymmetry in retail, wholesale, and shipping point pricing for fresh vegetables. *American journal of agricultural economics*, *64*(2), 205–212.

West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica: Journal of the Econometric Society*, *64*(5), 1067–1084.

White, H. (2000). A reality check for data snooping. *Econometrica*, *68*(5), 1097–1126.

Yang, H. & Ye, L. (2008). Search with learning: understanding asymmetric price adjustments. *The RAND Journal of Economics*, *39*(2), 547–564.

Zachariasse, V. & Bunte, F. (2003). How are farmers faring in the changing balance of power along the food chain. In *Beitrag zur Conference on Changing Dimensions of the Food Economy: Exploring the Policy Issues, Den Haag*, number 7, (pp. 1–11).