University of Copenhagen

ECONOMETRICS II

EXAM

Number: 9, 19 & 20

INDIVIDUAL CONTRIBUTIONS:

No. 9:
Sections 1.2, 1.4.3, 1.4.4, 1.4.5, 2.2, 2.5, 3.1.2, 3.1.4, 3.2.3

No. 19:
Sections 1.1, 1.4.1, 1.4.2, 1.5.1, 2.1, 2.4.2, 3.1.1, 3.2

No. 20:
Sections 1.3, 1.5, 2.2.4, 2.3, 2.4, 2.5.1, 3.1.3, 3.2.3

COUNT OF CHARACTERS:

Assignment 2: Approx. 12.100.

Assignment 4: Approx. 11.200.

New Assignment: Approx. 17.300.

May 27th 2019

Assignment 2

”Revisiting the Granger causality between stock prices and economic growth”

1.1 Introduction

This paper focuses on the dynamic relationship between GDP growth and real stock market returns. We examine, through a Granger-causality analysis, whether the stock market is a good predictor of future economic growth or economic growth drives stock market prices.

Feedback: Opgaven fik 100/100 point. Det blev bemærket, at afsnit 1.2 og 1.3 var en anelse langt.

Traditionally stocks have been viewed as a leading component for economic growth. One argument for this is that stock prices directly depend on expected profits and expected profits depend on the expectation of future economic growth. However, one may also argue for the opposite causality. The "wealth effect" indicate that if stock prices increase, private wealth increases, which in turn increases private consumption and economic growth.

To analyse the relationship we use a vector autoregression model (VAR). Our initial model is a VAR(7) which we revise to a VAR(2) by using the general-to-specific method. For the VAR(2) model we complete a Granger-causality test and compare results with relevant literature. The results of the test suggest that the stock market 'granger-causes' GDP growth, but reverse causality is not found. The conclusion is the same for both real and nominal returns.

1.2 Economic Theory

1.2.1 The Model

We consider a general vector autoregressive model of order k, in the following denoted VAR(k)-model, for the two-dimensional vector $Z_t = (\Delta Y_t, \Delta S_t)'$:

$$\begin{pmatrix}\Delta Y_t \\ \Delta S_t\end{pmatrix} = \begin{pmatrix}\mu_1 \\ \mu_2\end{pmatrix} + \begin{pmatrix}\Pi_{11}^1 & \Pi_{12}^1 \\ \Pi_{21}^1 & \Pi_{22}^1\end{pmatrix}\begin{pmatrix}\Delta Y_{t-1} \\ \Delta S_{t-1}\end{pmatrix} + \begin{pmatrix}\Pi_{11}^2 & \Pi_{12}^2 \\ \Pi_{21}^2 & \Pi_{22}^2\end{pmatrix}\begin{pmatrix}\Delta Y_{t-2} \\ \Delta S_{t-2}\end{pmatrix}$$
$$+ \ldots + \begin{pmatrix}\Pi_{11}^k & \Pi_{12}^k \\ \Pi_{21}^k & \Pi_{22}^k\end{pmatrix}\begin{pmatrix}\Delta Y_{t-k} \\ \Delta S_{t-k}\end{pmatrix} + \begin{pmatrix}\epsilon_{1t} \\ \epsilon_{2t}\end{pmatrix}, \quad t = 1,2,\ldots T \tag{1.2.1}$$

conditional on $Z_0, Z_{-1}, \ldots, Z_{-(k-1)}$. We assume errors conditioned on past information $\epsilon_t \mid \mathcal{I}_{t-1}$ are identically and individually distributed with a mean of 0 and covariance matrix given by Ω where $\mathcal{I}_{t-1} = \{Z_{t-1}, Z_{t-2}, \ldots, Z_{-(t-k)}\}$. We note that error terms ϵ_{1t} and ϵ_{2t} may be correlated.

Contemporanous effects Contemporanous effects are not directly parametized in the model, however causality between variables will be reflected by the error covariance Ω_{12} in the covariance matrix:

$$\Omega = \begin{pmatrix}\Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22}\end{pmatrix}, \tag{1.2.2}$$

where $\Omega_{12} = \Omega_{21}$ by symmetry.

Stationarity condition

The VAR(k)-model is stable and Z_t is stationary and weakly dependent if the eigenvalues of the *companion matrix*, given as:

$$\begin{pmatrix} \Pi_1 & \Pi_2 & \Pi_3 & \dots & \Pi_k \\ I_p & 0 & 0 & \dots & 0 \\ 0 & I_p & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & I_p & 0 \end{pmatrix} \tag{1.2.3}$$

are less than one in absolute value. We have here used that any VAR(k)-model can be written as a VAR(1) model on the *companion form*, cf. section 2.2 in Nielsen (2019a).

If the model is stationary and weakly dependent we find that Maximum Likelihood Estimators (MLEs) are consistent and normally distributed. The test statistics will have standard normal and χ^2-distributions, cf. section 4.1 in Nielsen (2019a).

Maximum Likelihood Estimation

Let $\theta = (\mu, \Pi^i, \Omega)'$ denote the parameters. Then the MLE is given as

$$\hat{\theta}(Z_1, \dots, Z_T) = \arg\max_{\theta} \log L(\theta \mid Z_1, \dots Z_T), \tag{1.2.4}$$

Of relevance here, we do not assume normality of errors. Rather, we noted that $\epsilon_t \mid \mathcal{I}_{t-1} \sim$ i.i.d(0, Ω) which is the requirement for MLEs to be asymptotically normally distributed, allowing for standard inference. We do so as tests for normality of errors are rejected, see section 1.4.2. Instead we estimate using Quasi-Maximum Likelihood (QMLE), which no longer ensures efficiency opposed to MLE, however still produces consistent estimates under stationarity. Note further that in the linear case the MLE coincides with OLS (Nielsen, 2019a, p.13).

1.2.2 Granger Causality

Note that the VAR model has no *a priori* assumptions on the causal direction between variables. However, we may examine causation with the notion of 'granger-causality'.

ΔS_t is said to 'granger-cause' ΔY_t if the expectation of the squared errors from forecasting is smaller when conditioned on both lagged values of ΔS_t and ΔY_t, rather than only ΔY_t. Note that this definition of causality relies on the cause to be previous to the effect.

Testing Granger-causality We can test the hypothesis of no granger-causality by imposing a restriction of the relevant variable. For example we may test the null $\Delta S_t \nrightarrow \Delta Y_t$ by imposing the restriction in eq. (1.2.1):

$$\Pi_{12}^i = 0 \text{ for } i = 1,2,\dots,k$$

If $k = 1$ we may use a simple t-test. For $k > 1$ one may use a Wald-test or a LR-test, with $\chi^2(k)$ distribution (or F-distribution in small samples). If we reject the null we say that ΔS_t 'granger-causes' ΔY_t.

1.2.3 Tests for misspecification

We test for misspecification using tests listed in table 1.1. Most importantly the model may not suffer from autocorrelation, as this imply that estimates will be inconsistent. Normality ensures that estimates converge quickly to the true value. Heteroscedasticity violates our stationarity assumption of constant variance, however may be circumvented by using robust errors.

Feedback: Man ville umiddelbart foretrække en LM-test fremfor en Portmanteau-test på grund af præcision. Portmanteau benyttes primært ved ikke-lineære modeller.

Table 1.1: Employed misspecification tests

Applied Test:	Null:	Test-statistic
Portmanteau(s)	No autocorrelation to lag s	χ^2
Normality	No significant deviation from normally distributed error terms	χ^2
Hetero-X	No heteroscedasticity (including cross-terms)	$\chi^2(2k)$

Source: Doornik and Hendry (2018a)
Notes: For degrees of freedom and details on small sample corrections, see Doornik and Hendry (2018a)

Test for lag determination

In section 1.4 we use a **general-to-specific** approach starting from 7 lags and reduce the lags by one until the last included lag is significant. Opposed to the specific-to-general approach all effects (including sufficient lags) likely to be relevant are included, then to have their validity tested (Doornik and Hendry, 2018b, pp. 153-154).

In order to determine the number of lags we use LR-testing to identify if imposed restriction are acceptable. If the null hypothesis given by imposed restriction is true, the test will be distributed:

$$LR(k = j - \frac{i}{p} \mid k = j) \overset{\mathrm{d}}{=} \chi^2(i), \tag{1.2.5}$$

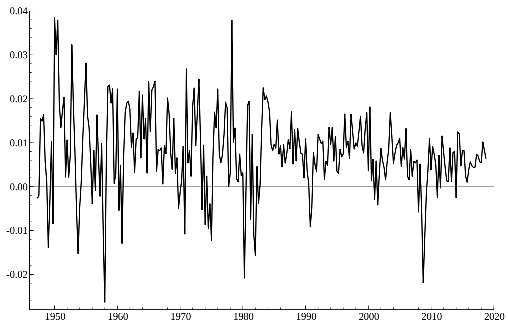
where j is the number of lags in the unrestricted model, i is the number of imposed restriction and p is the number of dimensions in the vector Z_t (here 2).

1.3 Description of Data

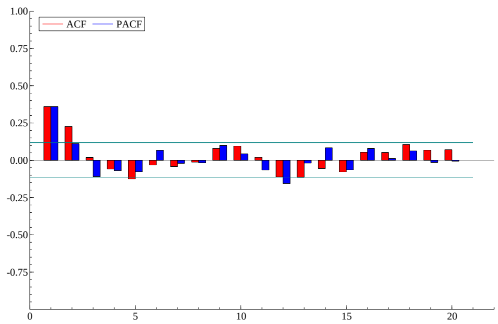
We define the variables:

- $Y_t = \log(\mathsf{GDP}_t)$ and $\Delta Y_t = Y_t - Y_{t-1}$, where GDP_t is the gross domestic product in billions of chained 2012 dollars, seasonally adjusted.
- $S_t = \log(\mathsf{SP500}_t/\mathsf{GDPdef}_t)$ and $\Delta S_t = S_t - S_{t-1}$, where $\mathsf{SP500}_t$ is the Standard & Poor's stock market index and GDPdef_t is the implicit price deflator for GDP, 2012=100, seasonally adjusted.
- $SN_t = \log(\mathsf{SP500}_t)$ and $\Delta SN_t = SN_t - SN_{t-1}$, where SN_t is corresponding nominal return of S_t

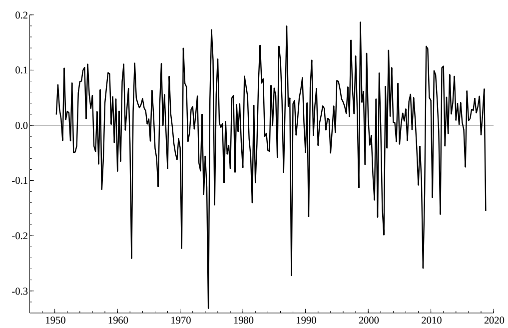
Figure 1.1: Autocorrelation functions and graphs for ΔY_t, ΔS_t and ΔSN_t



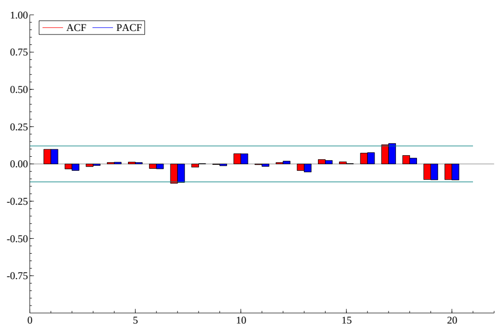
(a) Change in GDP (ΔY_t)



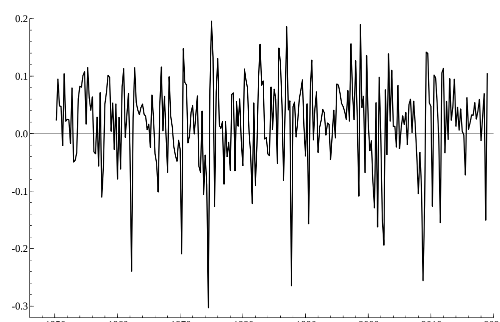
(b) PACF and ACF for ΔY_t



(c) Change in Deflated Stock Prices (ΔS_t)



(d) PACF and ACF for ΔS_t



(e) Change in nom. Stock Prices (ΔSN_t)

(f) PACF and ACF for ΔSN_t

Data is quarterly and is covering the period from 1947(1) to 2019(1) adopted from the FRED database maintained by the Federal Reserve Bank of St. Louis.

We take first differences of variables to ensure stationarity, see figure 1.1. To identify how many lags we want to describe our model with, we apply a Box-Jenkins identification approach by examining the partial autocorrelation function (PACF) for all three time series, see figure 1.1. We find that only the two first lags for ΔY_t are significant, which suggest choosing a VAR(2) model. To enable comparison with Comincioli (1996) and Foresti (2006), we initially look at a VAR(7) model and test for insignificant terms.

1.4 Empirical Model

1.4.1 Model selection

We test for heteroskedasticity, normality and no-autocorrelation. The models VAR(2)-VAR(7) do not suffer from heteroskedasticity nor autocorrelation, however we find that the residuals are not normally distributed, hence we use QMLE. Based on the misspecification test we discard VAR(1) and continue by looking at VAR(2)-VAR(7). We now compare all the preferred models and find the VAR model with 2 lags as being the best specified model according to the lowest AIC, HQ and SC values. Furthermore, we want to test the models VAR(3)-VAR(7) against VAR(2) by using the general-to-specific method. As all the models are nested we may conduct likelihood ratio tests on individual model-pairs. The tests are performed as described in the section 1.2.3, and show that we may easily accept reduction to the VAR(2) model.

1.4.2 Estimation results

We estimate the VAR(2)-model with ΔS_t and ΔSN_t for 273 observations, equations (1.4.2) and (1.4.3) with standard errors in parentheses. Further, see table 1.2 for misspecification test results.

Table 1.2: Misspecification test results

Applied Test	**Real model (1.4.2)**	**Nominal model (1.4.3)**
Portmanteau(12), $\chi^2(40)$	42.86 [0.3496]	42.79 [0.3523]
Normality, $\chi^2(4)$	65.77 [0.0000]	66.40 [0.0000]
Hetero-X, $F(42{,}760)$	1.097 [0.3138]	1.036 [0.4106]

Notes: Table of test-statistics. See table 1.1 for null hypothesis. P-values in brackets.

$$\begin{pmatrix}\Delta Y_t\\ \Delta S_t\end{pmatrix} = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} + \begin{pmatrix}\underset{(0.0584)}{0.2255} & \underset{(0.0062)}{0.02256}\\ \underset{(0.5805)}{0.3180} & \underset{(0.0617)}{0.0998}\end{pmatrix}\begin{pmatrix}\Delta Y_{t-1}\\ \Delta S_{t-1}\end{pmatrix} + \tag{1.4.1}$$
$$\begin{pmatrix}\underset{(0.0561)}{0.1038} & \underset{(0.0064)}{0.0260}\\ \underset{(0.5575)}{-0.8550} & \underset{(0.0632)}{-0.0411}\end{pmatrix}\begin{pmatrix}\Delta Y_{t-2}\\ \Delta S_{t-2}\end{pmatrix} + \begin{pmatrix}\epsilon_{1t}\\ \epsilon_{2t}\end{pmatrix}, \quad t = 1{,}2,\dots T$$

Nominal values ΔSN_t

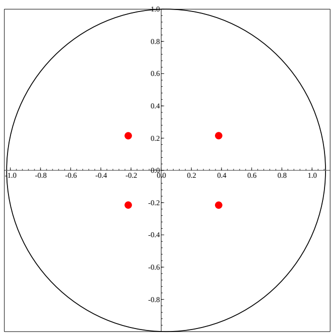
$$\begin{pmatrix} \Delta Y_t \\ \Delta SN_t \end{pmatrix} = \begin{pmatrix} \tilde{\mu}_1 \\ \tilde{\mu}_2 \end{pmatrix} + \begin{pmatrix} \underset{(0.0584)}{0.2263} & \underset{(0.0063)}{0.0230} \\ \underset{(0.5743)}{0.3385} & \underset{(0.0616)}{0.0861} \end{pmatrix} \begin{pmatrix} \Delta Y_{t-1} \\ \Delta SN_{t-1} \end{pmatrix} + \tag{1.4.2}$$

$$\begin{pmatrix} \underset{(0.0551)}{0.1046} & \underset{(0.0064)}{0.0264} \\ \underset{(0.5516)}{-0.8317} & \underset{(0.0631)}{-0.0553} \end{pmatrix} \begin{pmatrix} \Delta Y_{t-2} \\ \Delta SN_{t-2} \end{pmatrix} + \begin{pmatrix} \tilde{\epsilon}_{1t} \\ \tilde{\epsilon}_{2t} \end{pmatrix}, \quad t = 1,2, \dots T$$

Stationarity

From figure 1.2 and table 1.2 we see that the roots of the companion matrix are within the unit circle and the eigenvalues are less than one in absolute value.

Figure 1.2: Plotted roots of companion matrix (left)



Eigenvalues of companion matrix:		
real	imaginary	modulus
0.3811	0.2154	0.4377
0.3811	-0.2154	0.4377
-0.2184	0.2154	0.3065
-0.2184	-0.2154	0.3065

Table 1.3: Eigenvalues of companion matrix (right)

1.4.3 Recursive estimation

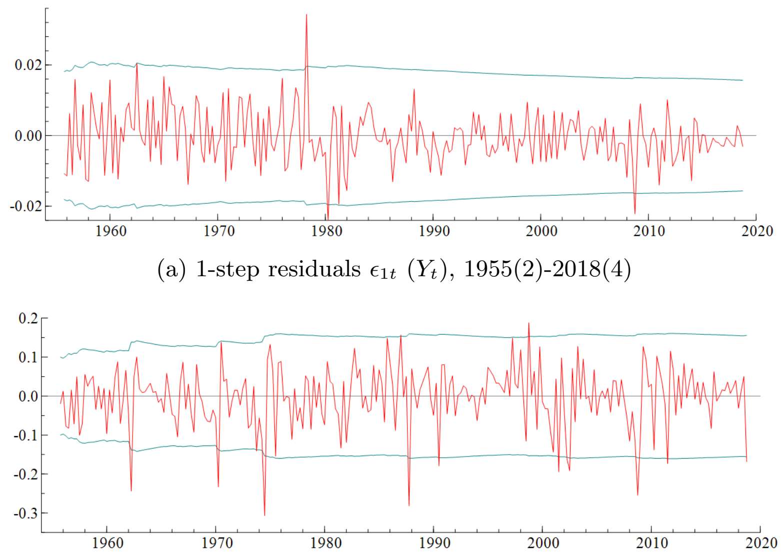
We may use recursive estimation to investigate behaviour of residuals. For examination we begin with a sample of 20 quarters and continue by increasing the sample by an additional quarter until we have used all observations. We calculate residuals as (Doornik and Hendry, 2018b, p. 47):

$$\epsilon_t = Z_t - \Pi^1 Z_{t-1} - \Pi^2 Z_{t-2} \tag{1.4.3}$$

Figure 1.3b and 1.3a show the time series for ϵ_{1t} and ϵ_{2t} plotted with bands for ± 2se. We observe a couple of outliers outside the confidence bands, which can indicate that the model has difficulties by handling outliers. It is expected that S_t have more and larger outliers than Y_t, as financial time series typically follows distributions with kurtosis larger than standard normal.

We might as such consider whether the model could be improved by using dummies for the relevant years or using a different error distribution.

Figure 1.3: Recursive estimation



(a) 1-step residuals ϵ_{1t} (Y_t), 1955(2)-2018(4)

(b) 1-step residuals ϵ_{2t} (S_t), 1955(2)-2018(4)

1.4.4 Test for Granger Causality

We examine the Granger-causality effect between stock prices and economic growth. We use a LR-test a $\chi^2(2)$-distribution. On a five-percent level the critical value is 5.99.

We formulate a VAR(2) model with no restrictions and note the log-likelihood estimate. Afterwards we test the following null-hypotheses of no-Granger causality with the listed restrictions;
The stock prices do not Granger-cause economic growth:

$$\Delta S_t \nrightarrow \Delta Y_t : \Pi_{12}^1 = \Pi_{12}^2 = 0 \tag{1.4.4}$$

The economic growth does not Granger-cause stock prices:

$$\Delta Y_t \nrightarrow \Delta S_t : \Pi_{21}^1 = \Pi_{21}^2 = 0 \tag{1.4.5}$$

It is possible to conclude from the test results that stock prices Granger-cause the economic growth (stars indicating rejection of null):

$$LR(\Delta S_t \nrightarrow \Delta Y_t) = 28.987^{**} \text{ and } LR(\Delta Y_t \nrightarrow \Delta S_t) = 2.1824 \tag{1.4.6}$$

Results are robust to using nominal values ΔSN_t instead of ΔS_t:

$$LR(\Delta SN_t \nrightarrow \Delta Y_t) = 29.693^{**} \text{ and } LR(\Delta Y_t \nrightarrow \Delta SN_t) = 2.197 \tag{1.4.7}$$

Feedback: Man kunne videreudvikle analysen med graf-teori – det ligger dog udenfor pensum.

1.4.5 Contemporaneous Effects

From our estimation we find that $\hat{\rho} = \text{corr}(\epsilon_{1t},\epsilon_{2t}) = 0.13761$. We wish to test whether this is significant, i.e. testing $\rho_0 = 0$. The standard variance for sample size T is given as

$\mathsf{se}(\hat{\rho}) = \frac{1}{\sqrt{T}}$. The test-statistic is given as:

$$t_{\rho=0} = \frac{\hat{\rho}}{\mathsf{se}(\hat{\rho})} = \frac{0.13671}{1\Big/\sqrt{273}} = 2.2737 \tag{1.4.8}$$

We may assume $\hat{\rho}$ to be asymptotically normal, in which case the calculated t-statistic is standard normal. We find that $t_{\rho=0}$ is larger than the critical value at a five-percent level and we may conclude that there is a significant contemporaneous effect between ΔY_t and ΔS_t. The conclusion is robust to using nominal values ΔSN_t.

1.5 Discussion and Concluding Notes

What might explain the Granger-causality? We find evidence that changes in stock prices ΔS_t is a predictor of economic growth. In the introduction we suggested this could be due to a 'wealth effect'. We find no evidence for a Granger-causal connection in the other direction.

Contemporaneous effects The definition of granger-causality (see section 1.2.2) assumes that the causing factor precedes the effect. This prevents us from defining any causal connection within the period. From section 1.4.5 we see that there is some contemporaneous effect within the period.

It cannot be rejected that within the same time period there might be some causal connection from $\Delta Y_t \rightarrow \Delta S_t$. We might suspect that any shocks to the real economy may be reflected in the stock prices within the time period and this to be explanatory for the found of contemporaneous effects. For instance erratic movements in economic growth due to external factors, such as economic crisis, political unrest, natural disasters, wars, etc. might affect stock prices, implying causality in the opposite direction within the time period. This point is left for future examination.

Outliers From the recursive estimation we observe several spikes outside the confidence band. For ΔS_t, where spikes are particularly prevalent, these reflect large and rapid changes in stock-prices. This is not unexpected as it is known that stock prices have larger kurtosis than the normal distribution. To improve the estimate it could be convenient to e.g. add dummy-variables for the involving years.

1.5.1 Other results

Comincioli (1996) makes the same findings as us; that stock prices Granger-cause the economic growth. In contrast, however, they found that the statistically significant lag length was three quarters.

Further, we reach the same conclusion as Foresti (2006), however we do not find significant lags up to order 7. Foresti's sample does not include the recent financial crisis – one event where we might suggest the found connection between stock prices and economic growth to be lacking. This, and methodological considerations, might explain differences in results.

Assignment 4

”Macroeconomics News Announcements and International Exchange Rates”

2.1 Introduction

This paper investigates the macroeconomic news announcements' effects on the exchange rate between the Euro (EUR) and the US Dollar (USD) applying an EGARCHM(1,1)-model inspired and simplified from Kim (1999). We find that both the surprise announcements of the number of employed and Purchasing Managers' Index have significant affect and appreciate the USD against the EUR. Moreover, the number of unemployed has significant effect on the conditional volatility and in addition we find evidence of asymmetric effects on conditional variance from news announcements.

Feedback: Opgaven fik 100/100 point.

2.2 Economic Theory

2.2.1 GARCH-models

Financial time series are often characterised by clusters of high volatility and low volatility also known as *volatility clustering*. One way of accommodating the clustering is by applying **general autoregressive conditional heteroskedasticity (GARCH)** models, that allow the variance of the error term ϵ_t to depend on its history.

2.2.2 Econometric Model

We wish to compare results from Kim (1999) and thus apply the same framework, namely an EGARCHM(1,1)-model. The model is chosen based on the following considerations. First, GARCH(1,1)-models have been shown to be useful for modelling **exchange rate changes**.

Second, as we wish to test whether there exists **asymmetric effects** of news announcements on the conditional variance, we employ an EGARCH model, which has shown to be simple and effective at modelling asymmetric effects. We might have applied the Threshold GARCH model (or GJR), however for comparison and due to the number of variables, empirical work may be eased using the EGARCH-framework. The model is asymmetric for $\gamma \neq 0$; in particular if $\gamma < 0$ negative shocks generate larger volatility than positive shocks[1].

Feedback: Pas på med at indføre notation, før den er blevet præsenteret i ligninger, her mht. γ

We note that **non-negativity** of the conditional variance σ_t^2 is required. Solving for σ_t^2 in (2.2.1c) reveals that the EGARCH-model by definition resolves any issues regarding negativity of the conditional variance, as $\forall x \in \mathbb{R} : \exp(x) > 0$, allowing us to add variables without concerns regarding negativity of variance term.

Finally, we might suspect that foreign investors holding US Dollars might require a risk premium if volatility increases. We model this by including the term $\alpha_\sigma \sigma_t$, i.e. conducting a **mean-variance** analysis, in (2.2.1a). Thus the model of choice is an EGARCH(1,1) in mean.

[1] We will only consider the modelled values of γ rather than undertaking sign bias test.

Model Specification

The model is given as[2]

$$\Delta s_t = \mu + \sum_{j=\text{EMP}}^{\text{CCI}} \alpha_j \text{NEWS}_{jt} + \alpha_\sigma \sigma_t + \epsilon_t \tag{2.2.1a}$$

$$\epsilon_t = z_t \sigma_t, \quad z_t \sim \text{GED}(0{,}1,\kappa) \tag{2.2.1b}$$

$$\ln \sigma_t^2 = \bar{\omega} + \beta_\sigma \ln \sigma_{t-1}^2 + \gamma \frac{\epsilon_{t-1}}{\sigma_{t-1}} + \varrho \frac{|\epsilon_{t-1}|}{\sigma_{t-1}} + \sum_{j=\text{EMP}}^{\text{CCI}} \beta_j \text{NEWS}_{jt}, \tag{2.2.1c}$$

where $\Delta s_t = 100 \cdot (\ln S_t - \ln S_{t-1})$,[3] σ_t^2 is the conditional variance of daily exchange rate changes, and NEWS_{jt} are the variables described in section 2.3.

The latter are included in (2.2.1a) in order to examine whether news announcement depreciates/appreciates the USD against the EUR, and in (2.2.1c) to test the effect on the conditional volatility, i.e. whether new information creates uncertainty on FX markets.

Drivers of variance We note that in standard (G)ARCH models ϵ_t drives the variance. In the EGARCH-model rather than ϵ_t, lagged values of $z_t = \frac{\epsilon_t}{\sigma_t}$ drives the variance. This implies that if a large shock ϵ_t arrives in a period of large variance σ_t^2 the effect will be less than if it arrives in another period with smaller variance, and vice versa. This feature is further a reason to consider the EGARCH model fit for our purpose.

Errors

As evident from table 2.4 we see that data is leptokurtic, as it is common for financial data. Hence, we are wrong to assume that errors are normal. Rather, we assume z_t to follow a General Error Distribution (GED) with tail-parameter $\kappa > 0$. We treat κ as an unknown parameter and estimate it jointly with other parameters. We note that $\kappa = 2$ equals the normal distribution, and due to the leptokurtic nature of data, expect κ to be less than 2.[4]

Estimator The non-normal distribution of errors has consequences for the asymptotic variance of our estimator, and we thus conduct QMLE rather than MLE, see section 1.2.1.

[2] We've used the following notation:

$$\sum_{j=\text{EMP}}^{\text{CCI}} \alpha_j \text{NEWS}_{jt} = \alpha_{\text{EMP}} \text{EMP}_t + \alpha_{\text{UEMP}} \text{UEMP}_t + \alpha_{\text{CPI}} \text{CPI}_t + \alpha_{\text{PMI}} \text{PMI}_t + \alpha_{\text{CCI}} \text{CCI}_t$$

Opposed to Kim (1999) we do not subtract $\sqrt{\frac{2}{\pi}}$ which implies that we should not interpret on the value of $\bar{\omega}$

[3] The scaling with 100 is used to make the maximum likelihood estimation more stable and log-returns comparable to percentage returns.

[4] In OxMetrics we estimate $\nu = \ln(\frac{\kappa}{2})$ equivalent to $\kappa = 2 \cdot \exp \nu$. The transformation is useful as we can use outputted t-tests to test for significance.

2.2.3 Finding ARCH-effects

To test for ARCH effects in the model we use the Breusch-Pagan LM test for no-heteroskedasticity (Nielsen, 2019d). This is done by regressing the squared residuals upon lagged values of squared residuals. In section 2.4 we test for ARCH(1)-effects using the below auxillary regression (however we may have extended for p lags of squared residuals):

$$\hat{\epsilon}_t^2 = \gamma_0 + \gamma_1 \hat{\epsilon}_{t-1}^2 + \text{error}. \tag{2.2.2}$$

We state the null hypothesis as no ARCH(1)-effects:

$$H_0 : \gamma_1 = 0.$$

We use the LM statistic $\xi_{ARCH} = T \cdot R^2$ as our test value, where R^2 is from the auxiliary regression (2.2.2). ξ_{ARCH} will be $\chi^2(1)$-distributed under the null.

2.2.4 Misspecification tests

Above we describe a test for ϵ_t to be correlated with ϵ_{t-1}, however it should still hold that $E(\text{NEWS}_{jt}\epsilon_t) = 0$ as estimates otherwise will be inconsistent. We test this using a Portmanteau-test, which will we be χ^2 distributed under the null of no autocorrelation (Doornik and Hendry, 2018b).

Heteroscedasticity violates our stationarity assumption of constant variance, however this may be circumvented by using robust errors.

2.3 Description of Data

Data covers daily data for the period January 4, 1999, to February 16, 2018, taken from Bloomberg. We define the macroeconomic news variables:

EMP$_t$ Surprise term in number of employed

UEMP$_t$ Surprise term in number of unemployed

CPI$_t$ Surprise term in consumer price inflation

PMI$_t$ Surprise term in Purchasing Managers' Index

CCI$_t$ Surprise term in Consumer Confidence Index

It is assumed that financial markets are informationally efficient as in Kim (1999) implying that interest rate and exchange rate represent equilibrium based on market participants expectations. Thus, for each of the mentioned macroeconomic variables, it is only the unexpected part of the surprise term that affects the market, as the expected part is already priced in the market. Announcements are made once a week or once a month, and for the remaining days the surprise term is zero.

It is noticed that the EUR was introduced in 1999, and no difference in currency-regimes have been observed, so we have no grounds for subdividing the sample.

It is evident from table 2.4 (kurtosis > 3) and figure 2.4 that the distribution is leptokurtic, meaning that ΔS_t is more peaked than the normal distribution and the tails are "fatter" compared to the normal distribution. We further note that the mean is fairly close to zero, as expected.

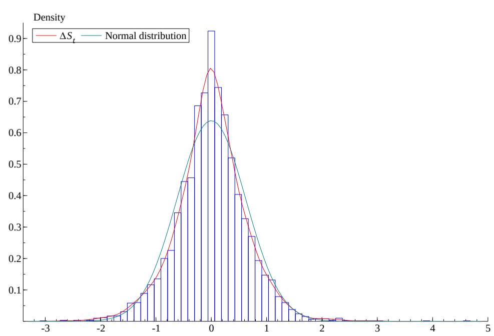
Feedback: Data vil som følge af GARCH-effekter have kurtosis over 3. Det har dog ikke påvirkning af fejlleddet og afsnittet er således overflødigt og tjener ikke det formål, der beskrives. Man kunne have lavet et Q-Q-plot af fejlled, der ville være meget lig figur 2.4, da der ingen struktur er i middelværdier.

Table 2.4: Descriptive Statistics for ΔS_t

	Mean	Variance	Skewness	Kurtosis
ΔS_t	0.0011	0.3903	0.1178	5.2202

Notes: Statistics are calculated using Descriptive Statistics package v. 1.8 for OxMetrics, Copyright of Heino Bohn Nielsen. We note that mean is close to 0 and kurtosis larger than 3.

Figure 2.4: Distribution of ΔS_t against the normal distribution



Notes: Graph included for illustrative purposes highlighting the larger kurtosis.

2.4 Empirical Results

2.4.1 Misspecification

In table 2.5 we report results from misspecification test on the EGARCHM(1,1)-model. We find evidence for ARCH-effects and no autocorrelation. As predicted we do not find normality of errors. We mention that the model is leptokurtic – further as seen in table 2.5 we find that $\ln(\nu/2)$ is significantly different from 0.

2.4.2 Estimation Results

Number of employed and unemployed

The news announcement surprise term of the number of employed has a significant effect on the mean, but no significant effect on the conditional volatility. When the news are released the US dollar appreciates against the EUR by 0.0031 percent (note the converse sign of estimates). Further, the news announcement in the number of unemployed has a

Table 2.5: The table shows estimates of the model in equation (2.2.1a) and equation (2.2.1c). Standard errors in $(\cdot)$ and p-values in $[\cdot]$ for misspecification tests.

	EGARCH-M(1,1)
μ (2.2.1a)	0.0396 (0.0314)
α_{CCI} (2.2.1a)	-0.0103 (0.0073)
α_{CPI} (2.2.1a)	0.131 (0.343)
α_{UEMP} (2.2.1a)	0.0000912 (0.0009)
α_{EMP} (2.2.1a)	-0.0031 (0.0007)
α_{PMI} (2.2.1a)	-0.0820 (0.0227)
$\sqrt{h_t}$ (2.2.1a)	-0.0659 (0.0586)
$\bar{\omega}$ (2.2.1c)	-0.0063 (0.0019)
β_{σ} (2.2.1c)	0.995 (0.0016)
β_{CCI} (2.2.1c)	-0.0106 (0.006)
β_{CPI}(2.2.1c)	0.0435 (0.170)
β_{UEMP} (2.2.1c)	0.0028 (0.0008)
β_{EMP} (2.2.1c)	-0.0005 (0.0003)
β_{PMI} (2.2.1c)	0.0119 (0.0140)
γ, asymmetry (2.2.1c)	-0.0149 (0.0051)
ϱ (2.2.1c)	0.0583 (0.0088)
$\ln(\nu/2)$ (GED)	-0.311 (0.0333)
Log-lik.	-4177.5
Portmanteau, 1-69	[0.95]
No ARCH(1) (ξ_{ARCH})	[0.02]
Normality	[0.00]
T	4812
Sample start	1999-01-05
Sample end	2018-02-16

Notes: Std. errors in parentheses. P-values in squares. Notice that α refers to equation 2.2.1a and β refers to equation 2.2.1c.

significant positive effect on the conditional volatility by 0.0028 percent, but not on the mean.

In theory an unexpected higher number of employed indicates an unexpected boost in the economy, which will appreciate the USD against the EUR Kim (1999).

Consumer price inflation

The CPI news term has no significant effect on either the mean nor the conditional volatility. However, in theory we would expect that if a positive surprise occurs the US dollar will appreciate against the EUR, because the central bank will respond to a rise in the inflation by raising the interest rate, which will slow down the economy but strengthen the USD against other currencies.

Purchasing Managers Index

The PMI surprise term has a significant effect on the mean, but no significant effect on the conditional volatility. When the PMI news announcement is released the USD appreciates against the EUR by 0.0820 percent. If a positive shock hits the manufacturing and service sectors and hereby has a positive effect on the business conditions it will indicate a boost to the economy. This will appreciate the US dollar against the EUR.

Consumer confidence Index

The CCI surprise term has no significant effect on the the conditional volatility nor the mean. When the CCI news announcement is released we would expect an increase in the CCI to be associated with a USD appreciation (David Gulley and Sultan, 1998).

Asymmetric effects

We find evidence for asymmetric effects of markets shocks on the conditional variance. We find that γ is significant and *negative*, indicating that negative shocks have larger impact on the conditional variance, than positive. We note that the notion of 'good' and 'bad' news, commonly used in analysis of stock prices, should not apply to exchange rates as agents are on both sides of the market.

Mean-Variance Analysis

We do not find evidence of appreciation of the US dollar following larger volatility. This is consistent with Kim (1999) and we might speculate that an effect would more likely be evident only in smaller currencies.

2.5 Discussion and concluding notes

The large coefficient on $\ln \sigma^2_{t-1}$, $\beta_\sigma = 0.995$ reflects high persistence in the volatility of the exchange rate. As such one might consider modelling using an integrated-GARCH approach where volatility shocks have permanent effects. However, we generally find that EGARCHM(1,1)-model framework with a General Error Distribution accounts for the observed properties in table 2.4 well.

Feedback: Det ville ikke være meningsfyldt at benytte en I-GARCH model, når koefficienterne ikke er tættere på nul; det ville medføre model med trend. Den høje β-værdi er bedre set som udtryk for høj persistens.

Lack of market response might be a sign of missing consensus in the market on e.g. policy responses. This might explain the missing response to surprises in CPI as the market response depends on the expected policy changes.

Kim (1999) finds increased volatility on days of announcements (for e.g. the AUDUSD exchange rate). This might be an interesting topic to follow in further research, however has not been tested here.

Overall, we find that the news announcement for the number of employed and PMI significantly affect the USD to appreciate against the EUR. Moreover we find that the news announcement for the number of unemployed significantly affects the conditional variance.

2.5.1 Comparison with Kim (1999)

Results are generally in line with Kim (1999) that finds the surprise term in the number of employed has a significant negative effect on the exchange rate. However, no significant effect from the CPI news is found, which also corresponds to our results. Alike Kim (1999) we conclude that asymmetry occurs in the conditional volatility, and find that negative news has a larger impact on the exchange rate than positive news.

New Assignment

"The Phillips Curve"

3.1 Theoretical Assignment

3.1.1 Measurement error (1a)

We examine an expectation augmented Phillips curve, given as:

Feedback: Opgaven fik 95/100 point. 5 point blev fratrukket for ikke at foreslå $\pi^{\text{survey}}_{t-1}$, men π_{t-1}

$$\pi_t = \alpha_1 + \alpha_2 u_t + \alpha_3 \pi^e_{t+1|t} + \eta_t, \quad t = 1,2,\ldots,T, \tag{3.1.1}$$

with a given measure of expected inflation:

$$\pi^{\text{survey}}_{t+1|t} = \pi^e_{t+1|t} + v_t, \tag{3.1.2}$$

where v_t is a measurement error uncorrelated with u_t and η_t. We wish to examine whether we can use OLS to estimate $\beta = (\alpha_1,\alpha_2,\alpha_3)$. We rewrite model (3.1.1) by using the information given in equation (3.1.2) and find:

$$\pi_t = \alpha_1 + \alpha_2 u_t + \alpha_3 \pi^{\text{survey}}_{t+1|t} + \epsilon_t, \tag{3.1.3}$$

where

$$\epsilon_t = \eta_t - \alpha_3 v_t$$

We note that the OLS estimator $\hat{\beta}$ is inconsistent as both ϵ_t and $\pi^{\text{survey}}_{t+1|t}$ depend on v_t. This implies that $E(\pi^{\text{survey}}_{t+1|t}\epsilon_t) \neq 0$, violation of a necessary condition for consistency.

Derivation We may illustrate this in deeper detail. For simplicity, set $\alpha_2 = 0$ (however Verbeek (2008) also notes that results holds for the multivariate case, too). We can write the estimator for α_3, where $\bar{\pi}^{\text{survey}}_{t+1|t}$ denotes the sample mean, as

$$\hat{\alpha}_3 = \frac{\sum_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})(\pi_t - \bar{\pi}_t)}{\sum_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})^2} \tag{3.1.4}$$

We rewrite using equation (3.1.3):

$$\begin{aligned}
\hat{\alpha}_3 &= \frac{\frac{1}{T}\Sigma_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})(\alpha_1 + \alpha_3 \pi^{\text{survey}}_{t+1|t} + \epsilon_t - \alpha_1 - \alpha_3 \bar{\pi}^{\text{survey}}_{t+1|t} - \bar{\epsilon}_t)}{\frac{1}{T}\Sigma_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})^2} \\
&= \frac{\frac{1}{T}\Sigma_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})(\alpha_3(\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t}) + (\epsilon_t - \bar{\epsilon}_t))}{\frac{1}{T}\Sigma_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})^2} \\
&= \frac{\alpha_3 \cdot \frac{1}{T}\Sigma_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})^2 + \frac{1}{T}\Sigma_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})(\epsilon_t - \bar{\epsilon}_t)}{\frac{1}{T}\Sigma_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})^2} \\
&= \alpha_3 + \frac{\frac{1}{T}\Sigma_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})(\epsilon_t - \bar{\epsilon}_t)}{\frac{1}{T}\Sigma_{t=1}^T (\pi^{\text{survey}}_{t+1|t} - \bar{\pi}^{\text{survey}}_{t+1|t})^2}
\end{aligned}$$

As the sample size tends to infinity sample moments converge to population moments.

$$\hat{\alpha}_3 \xrightarrow{\text{p}} \alpha_3 + \frac{E(\pi^{\text{survey}}_{t+1|t}\epsilon_t)}{V(\pi^{\text{survey}}_{t+1|t})} = \alpha_3 - \alpha_3 \frac{V(v_t)}{V(\pi^{\text{survey}}_{t+1|t})} \tag{3.1.5}$$

See section 3.2.3 in appendix for derivation of the latter equality. This shows that $\hat{\alpha}_3$ is only a consistent estimator if the variance of v_t is equal to zero, i.e. if the measurement error is non-existing.

Valid instruments

We wish to suggest a list of instruments and show that they are valid.

A variable can be used as a *valid* instrument for the explanatory variable $\pi_{t+1|t}^{\text{survey}}$ if it is **relevant** and **exogenous**. Relevant refers to the condition that the variable should be correlated with the endogenous variable $\pi_{t+1|t}^{\text{survey}}$: $\text{Cov}(z_t, \pi_{t+1|t}^{\text{survey}}) \neq 0$. Exogenous refers to the condition that the instrument variable should be uncorrelated with the error term: $\text{Cov}(z_t, \epsilon_t) = 0$.

We argue that lagged values of π_t (internal instrument) may be used to instrument $\pi_{t+1|t}^{\text{survey}}$. We may argue that this is exogenous:

$$\begin{aligned} \text{Cov}(\pi_{t-1}, \epsilon_t) &= \text{Cov}(\pi_{t-1}, \eta_t - \alpha_3 v_t) \\ &= \text{Cov}(\pi_{t-1}, \eta_t) - \alpha_3 \text{Cov}(\pi_{t-1}, v_t) = 0, \end{aligned} \tag{3.1.6}$$

where we note that since π_{t-1} is given in period t, we find $\text{Cov}(\pi_{t-1}, \eta_t) = \pi_{t-1} E(\eta_t) = 0$ and equally π_{t-1} with the measurement error, v_t. The relevancy of instruments may be tested empirically. It is commonly assessed that internal instruments are relevant for time-series (Nielsen, 2019b), which is also shown in the empirical exercise.

Trivially, the number 1 fulfills the relevancy and exogeneity criteria for a constant and may be used to instrument α_1. Further, u_t may instrument itself. It can likewise be argued to be exogenous, as: $\text{Cov}(u_t, \epsilon_t) = \text{Cov}(u_t, \eta_t - \alpha_3 v_t) = 0$ as we note that both η_t and v_t are uncorrelated with u_t as by definition.

Accordingly, we may suggest a vector of instruments:

$$z_t = \begin{pmatrix} 1 \\ u_t \\ \pi_{t-1} \end{pmatrix} \tag{3.1.7}$$

The above argumentation could be extended for other lags of π_t and u_t, however long lags may compromise the relevancy criteria.

Feedback: Her burde man foreslå $\pi_{t-1}^{\text{survey}}$ i stedet for π_{t-1}, da man må forestille sig, at $\pi_{t-1}^{\text{survey}}$ tættere korreleret med $\pi_{t+1|t}^{\text{survey}}$. Det er tilladt, da $v_t \sim$ i.i.d.

Sample moment conditions and MM estimator

To simplify notation we can rewrite equation (3.1.1) in terms of vectors:

$$\pi_t = x_t' \beta + \epsilon_t \tag{3.1.8}$$

Where $x_t = (1,\ u_t,\ \pi_{t+1|t}^{\text{survey}})$ denotes the model variables and $\beta = (\alpha_1,\ \alpha_2,\ \alpha_3)$ denotes a vector of the true value of the parameters. The sample moment conditions can be written as:

$$\begin{aligned} g_T(\hat{\beta}) &= \frac{1}{T}\sum_{t=1}^{T} z_t\epsilon_t \\ &= \frac{1}{T}\sum_{t=1}^{T} z_t(\pi_t - x'\hat{\beta}) \\ &= \frac{1}{T}\sum_{t=1}^{T} z_t(\pi_t - \alpha_1 - \alpha_2 u_t - \alpha_3\pi_{t+1|t}^{\text{survey}}) = 0 \end{aligned}$$

In the case of three moment conditions we have exact identification and we can derive the method of moment estimator $\hat{\beta}$ from the sample moment conditions:

$$\begin{aligned} g_T(\hat{\beta}_{MM}) &= \frac{1}{T}\sum_{t=1}^{T} z_t(\pi_t - x_t'\hat{\beta}_{MM}) = 0 \\ &\Leftrightarrow \frac{1}{T}\sum_{t=1}^{T} z_t\pi_t - \frac{1}{T}\sum_{t=1}^{T} z_t x_t'\hat{\beta}_{MM} = 0 \\ &\Leftrightarrow \frac{1}{T}\sum_{t=1}^{T} z_t\pi_t = \frac{1}{T}\sum_{t=1}^{T} z_t x_t'\hat{\beta}_{MM} \\ &\Leftrightarrow \hat{\beta}_{MM} = (\frac{1}{T}\sum_{t=1}^{T} z_t x_t')^{-1}\frac{1}{T}\sum_{t=1}^{T} z_t\pi_t \end{aligned}$$

We note that the matrix $\frac{1}{T}\sum_{t=1}^{T} z_t x_t'$ must be non-singular and thus invertible. Notice that the $\hat{\beta}_{MM}$ corresponds to the IV estimator.

3.1.2 Measurement error with known functional form (1b)

Assume now that v_t follows an MA(1) process, such that

$$v_t = \xi_t + \rho\xi_{t-1} \tag{3.1.9}$$

Substitute (3.1.9) with (3.1.2) into (3.1.1) and find

$$\pi_t = \alpha_1 + \alpha_2 u_t + \alpha_3\pi_{t+1|t}^{\text{survey}} + \epsilon_t, \tag{3.1.10}$$

where

$$\epsilon_t = \eta_t - \alpha_3(\xi_t + \rho\xi_{t-1})$$

Corresponding to section 3.1.1 we note that we cannot use OLS to estimate β, as the estimator is inconsistent because both ϵ_t and $\pi_{t+1|t}^{\text{survey}}$ depend on $v_t = \xi_t + \rho\xi_{t-1}$. This implies that conditions for consistency are violated; $E(\pi_{t+1|t}^{\text{survey}}\epsilon_t) \neq 0$. The same argumentation follows as above, see section 3.1.1

Feedback: Nu er v_t ikke længere i.i.d. og det kræver således et nyt (og mere spændende) svar end hvis man havde brugt π_{t-1}.

Valid instruments

For show, we might consider π_{t-1} an instrument as above:

$$\begin{aligned}\text{Cov}(\pi_{t-1},\epsilon_t) &= \text{Cov}(\pi_{t-1},\eta_t - \alpha_3(\xi_t + \rho\xi_{t-1})) \\ &= \text{Cov}(\pi_{t-1},\eta_t) - \alpha_3[\text{Cov}(\pi_{t-1},\xi_t) + \rho\text{Cov}(\pi_{t-1},\xi_{t-1})]\end{aligned} \tag{3.1.11}$$

We cannot be assured that $\text{Cov}(\pi_{t-1},\xi_{t-1}) = 0$ and thus π_{t-1} does not fulfil the exogeneity requirement. Interchanging with π_{t-2} we find that $\text{Cov}(\pi_{t-2},\epsilon_t) = 0$. Consequently, we suggest the following vector of instruments:

Feedback: $\text{Cov}(\pi_{t-1},\xi_{t-1})$ kan i princippet være 0, argumentet er tydeligere, når 'survey'-værdien anvendes, da den med sikkerhed er forskellig fra 0.

$$\tilde{z}_t = \begin{pmatrix} 1 \\ u_t \\ \pi_{t-2} \end{pmatrix} \tag{3.1.12}$$

In both cases above we might add additional lags of the endogenous variables as long as they fulfill the relevance criteria. One could further add external instruments to the list (e.g. inflation expectations derived from financial instruments), however rigorousness of argumentation on exogeneity may be compromised for external instruments.

Sample moment conditions

The sample moment conditions can be written as

$$\begin{aligned}g_T(\hat{\beta}) &= \frac{1}{T}\sum_{t=1}^{T} \tilde{z}_t \epsilon_t \\ &= \frac{1}{T}\sum_{t=1}^{T} \tilde{z}_t(\pi_t - x'\hat{\beta}) = 0\end{aligned}$$

3.1.3 Forward-looking model with rational expectations (2a)

We consider the following model (note that this section also serves as theoretical background for empirical estimation):

$$\pi_t = \alpha_1 + \alpha_2 u_t + \alpha_3 \pi^e_{t+1|t} + \eta_t, \quad t = 1,\ldots,T \ , \tag{3.1.13}$$

with rational (or model-consistent) expectations:

$$\pi^e_{t+1|t} = E(\pi_{t+1}|\mathcal{I}_t), \tag{3.1.14}$$

where

$$\mathcal{I}_t = \{\pi_{t-1},\pi_{t-2},\ldots,\pi_1; u_t,u_{t-1},\ldots,u_1\} \tag{3.1.15}$$

We cannot estimate (3.1.13) directly due to the latency of $E(\pi_{t+1} \mid \mathcal{I}_t)$ and thus replace it by π_{t+1}:

$$\pi_t = \alpha_1 + \alpha_2 u_t + \alpha_3 \pi_{t+1} + e_t, \tag{3.1.16}$$

where $e_t = \eta_t - \alpha_3 (\pi_{t+1} - E(\pi_{t+1} \mid I_t))$, such that $\pi_{t+1} - E(\pi_{t+1} \mid I_t)$ declares the forecast error in predicting future inflation. By construction the relation faces correlation between the error term e_t and explanatory variables, and we may thus not apply simple linear regression.

Rational expectations implies that all variables in the information set $\mathcal{I}_t$ are uninformative of forecasting errors $\pi_{t+1} - E(\pi_{t+1} \mid I_t)$, implying (as η_t is uncorrelated with $\pi^e_{t+1|t}$):

$$E(e_t \mid I_t) = 0, \tag{3.1.17}$$

which further implies that

$$E(e_t z_t) = 0, \tag{3.1.18}$$

for $z_t \in \mathcal{I}_t$.

Moment conditions The population moment conditions can thus be written as:

$$g(\beta_0) = E(z_t e_t) = E[z_t(\pi_t - \alpha_1 - \alpha_2 u_t - \alpha_3 \pi_{t+1})] = 0 \tag{3.1.19}$$

and corresponding sample moment condition

$$g(\hat{\beta}) = \frac{1}{T} \sum_{t=1}^{T} z_t(\pi_t - \alpha_1 - \alpha_2 u_t - \alpha_3 \pi_{t+1}) = 0 \tag{3.1.20}$$

which may be rewritten on matrix notation as:

$$g(\hat{\beta}) = T^{-1} Z'(\Pi - X\beta) = 0 \tag{3.1.21}$$

Choice of estimator

Due to correlation between e_t and explanatory variables, we may not use OLS. Further MLE would require full information of distributions. Here we have more moment conditions than parameters $R > K$, as we may include multiple instruments, hence we should use General Method of Moments (GMM) rather than Method of Moments (of which IV is a special case).

GMM estimator With regard to GMM we might consider different estimation strategies in order to obtain the optimal weight matrix and the efficient GMM estimator.

We note that the influential paper by Galí and Gertler (1999) on the New Keynesian Phillips Curve uses a two-step GMM estimator. However, using iterated GMM estimation, where the parameters and weight matrix are interchangeably estimated until convergence, has shown to perform better in certain samples (note that the two-step approach is asymptotically equivalent to the iterated estimator).

Economic literature has not proven the continuously updated estimator to be superior to iterated GMM in practice. In the empirical part, estimation results (see table 3.6) are included for both two-step GMM estimation (for similarity with Galí and Gertler (1999)) and iterated GMM estimation.

Choice of weights

The GMM estimator includes a weight matrix (see derivation below, section 3.1.4) to ensure that no single moment overrules the minimisation. Weights are assigned according to precision, that is low variance moments are assigned higher weights than high variance moments. It must hold that the weight matrix is positive definite and puts a non-zero weight on all moment conditions.

For consistency of the estimator we require data to be stationary and weakly dependent. GMM estimators vary with different weights, however for a given data set, they will all be consistent.

We wish to estimate the optimal weight matrix, which is given as $W_T^{\text{opt}} = S_T^{-1}$, where S_T is a consistent estimator of $S = T \cdot V(g_T(\theta_0))$. How to construct the estimator depends on the properties of data:

Independent and Identically Distributed (IDD): Used if the data are independent.

Heteroscedasticity Consistent (HC): Robust to heteroscedasticity.

Heteroskedasticity and Autocorrelation Consistent (HAC): Using HAC, we are allowing heteroskedasticity, and furthermore we are allowing autocorrelation. Contributions from covariances are included as data are no longer independent. We cannot consistently estimate for as many covariances as we have observations ($\hat{S_T}$ is not positive definite with probability 1) and thus employ kernel estimators, which allow weights to go towards zero. Here we choose Bartlett Kernel weights; these ensure weights decrease linearly. For further technical details on weights, we refer to (Nielsen, 2019b, sec. 5).

We might suspect errors to suffer from both heteroscedasticity and serial correlation, and as a result refer to HAC(12)-estimator. However, as Mavroeidis (2005) has pointed to, the HAC-estimator compromises the power of the J-tests, see section 3.2.2.

3.1.4 Deriving GMM estimator (2b)

The GMM estimator is the argument that minimises the criteria function $Q_T(\beta)$:

$$\hat{\beta}_{GMM} = \arg\min_{\beta}\{Q_T(\beta)\} \tag{3.1.22}$$

We write the criteria function using previously introduced matrix notation:

$$\begin{aligned} Q_T(\beta) &= g_T(\beta)' W_T g_T(\beta) \\ &= (T^{-1}Z'(\Pi - X\beta))' W_T (T^{-1}Z'(\Pi - X\beta)) \\ &= (T^{-1}(\Pi' - \beta'X')Z) W_T (T^{-1}Z'(\Pi - X\beta)) \\ &= T^{-2}(\Pi' - \beta'X')ZW_TZ'(\Pi - X\beta) \end{aligned}$$

Expand all terms:

$$= T^{-2}\left(\Pi'ZW_TZ'\Pi - \Pi'ZW_TZ'X\beta - \beta'X'ZW_TZ'\Pi - \beta'X'ZW_TZ'X\beta\right)$$

Use that $\Pi'ZW_TZ'X\beta$ is a scalar variable, cf. (Nielsen, 2019c, p.6), such that $\Pi'ZW_TZ'X\beta = (\Pi'ZW_TZ'X\beta)' = \beta'X'ZW_TZ'\Pi$:

$$= T^{-2}\left(\Pi'ZW_TZ'\Pi - 2\beta'X'ZW_TZ'\Pi - \beta'X'ZW_TZ'X\beta\right)$$

We take the first derivative, applying results from Nielsen (2019c):

$$\begin{aligned}\frac{\partial Q_T(\beta)}{\partial\beta} &= T^{-2}(-2X'ZW_TZ'\Pi + (X'ZW_TZ'X + X'ZW_TZ'X)\beta \\ &= T^{-2}(-2X'ZW_TTZ'\Pi + 2X'ZW_TZ'X\beta) \\ &= 2T^{-2}X'ZW_TZ'X\beta - 2T^{-2}X'ZW_TZ'\Pi\end{aligned}$$

We solve:

$$\begin{aligned}\frac{\partial Q_T(\hat{\beta}_{GMM})}{\partial\hat{\beta}_{GMM}} &= 0 \\ 2T^{-2}X'ZW_TZ'X\hat{\beta}_{GMM} &= 2T^{-2}X'ZW_TZ'\Pi\end{aligned}$$

Provided that $X'ZW_TZ'X$ is non-singular and invertible (this ensures no perfect multicollinearity):

$$\hat{\beta}_{GMM}(W_T) = (X'ZW_TZ'X)^{-1}X'ZW_TZ'\Pi$$

Second derivative To ensure that we found a minimum and not a maximum we take the second derivative:

$$\frac{\partial^2}{\partial\beta\partial\beta'}Q_T(\beta) = \frac{\partial}{\partial\beta'}(2T^{-2}X'ZW_TZ'X\beta - 2T^{-2}X'ZW_TZ'\Pi)$$

Apply rule (6$\star$) in Nielsen (2019c):

$$= 2T^{-2}X'ZW_TZ'X$$

We note that as W_T is positive definite by definition the second derivative is positive definite by construction, indicating that $\hat{\beta}_{GMM}(W_T)$ is a minimum.

3.2 Empirical Assignment

3.2.1 Estimation of model (3.1.13) with expectations (3.1.14)

We estimate the model using four lagged values of inflation π_t and output gap u_t (measured as unemployments deviation from mean), i.e. $z_t = (1; u_t, u_{t-1}, \ldots, u_{t-4}; \pi_{t-1}, \ldots, \pi_{t-4})$. Attempts using more lagged values have been conducted, however, test on incremental J-statistics (Nielsen, 2019b, p. 19) have shown that fewer lags are preferable.

For derivation of sample moment conditions and considerations on applied estimator and weights, see section 3.1.3.

Estimation results

For estimations of model (3.1.16) we find that all estimates for both π_{t+1} and u_t are positive and significant at a 5 pct. significance level with values around 1.0 and 0.06 respectively. We note that a value of unity of α_3 is in line with results from Galí and Gertler (1999). Measures of output gap are slightly different – however note that Galí and Gertler (1999) uses marginal costs to proxy output gap.

Feedback: Der mangler måske en kort forklaring af skiftende fortegn på estimationsresultater.

Estimators Results are reported for both iterated and two-step GMM in table 3.6. We find that there is slight differences between the estimation methods (primarily when applying HAC weights), however small deviations are in line with the observation that the two methods are asymptotically equivalent, see section 3.1.3. We therefore allow for comparison with e.g. Galí and Gertler (1999) who applies a two-step estimation.

Table 3.6: Estimation results for the forward-looking model and hybrid model

	Estimator	**Weight**	**Iter.**	α_1	α_2	α_3	α_4	Q_T	ξ_J	**DF**
F	Iterated GMM	IID	4	0.002020	0.07117	0.9999	-	0.076887	31.60	7
				(0.03273)	(0.0128)	(0.005825)			[0.000]	
F	Iterated GMM	HC	10	0.02974	0.06861	0.9934	-	0.049739	20.44	7
				(0.03137)	(0.01405)	(0.006698)			[0.005]	
F	Iterated GMM	HAC	29	0.05053	0.07050	0.9943	-	0.026579	10.92	7
				(0.04697)	(0.01891)	(0.01061)			[0.142]	
F	Two-step	IID	2	0.002020	0.07117	0.9999	-	0.075790	31.15	7
				(0.03296)	(0.01289)	(0.005867)			[0.000]	
F	Two-step	HC	2	0.02293	0.06767	0.9951	-	0.051684	21.24	7
				(0.03205)	(0.01414)	(0.006959)			[0.003]	
F	Two-step	HAC	2	0.01645	0.05951	1.00062	-	0.028144	11.57	7
				(0.04731)	(0.01906)	(0.01104)			[0.116]	
H	Iterated GMM	IID	4	−0.01348	0.004762	0.5573	0.4455	0.026447	10.87	6
				(0.01984)	(0.01086)	(0.05095)	(0.05117)		[0.092]	
H	Iterated GMM	HC	9	−0.01008	0.006104	0.5530	0.4480	0.025819	10.61	6
				(0.01953)	(0.01131)	(0.05507)	(0.05622)		[0.101]	
H	Iterated GMM	HAC	30	−0.01028	−0.003205	0.4934	0.5084	0.025736	10.58	6
				(0.007776)	(0.005219)	(0.02974)	(0.02992)		[0.0257]	
H	Two-step	IID	2	−0.01348	0.004762	0.5573	0.4455	0.025612	10.53	6
				(0.02016)	(0.01103)	(0.05177)	(0.05199)		[0.104]	
H	Two-step	HC	2	−0.009823	0.0064395	0.5555	0.4453	0.025347	10.42	6
				(0.01977)	(0.01144)	(0.05612)	(0.05717)		[0.108]	
H	Two-step	HAC	2	−0.007547	0.00210962	0.5397	0.4613	0.023706	9.74	6
				(0.01135)	(0.006292)	(0.03443)	(0.03407)		[0.136]	

Notes: Std. errors in parentheses. p-values in squares. IID: Independent and identically distributed. HC: Allowing for heteroskedasticity of the moments. HAC: allowing for heteroscedasticity and autocorrelation. Q_T: Criteria Function. ξ_J is J-test statistic. F: forward-looking model. H: Hybrid model. All regressions are estimated with $T = 411$.

3.2.2 Overidentifying restrictions

We use J-test (or Hansen test) for testing over-identifying restrictions on the moment conditions. Under the null that the over-identifying restrictions are valid, the statistic is distributed as a χ^2 with $R - K$ degrees of freedom, where R is the number of moment conditions and K is the number of parameters.

The test examines whether the overidentifying $R-K$ restriction are correct. If restrictions are far from zero it may indicate that the moment conditions are violated Nielsen (2019b).

Interpretation and concerns

Results of J-tests (ξ_J) are reported in table 3.6. We find that we reject the J-tests when using IID and HC weights. This might imply that some of the moment conditions are violated.

On the other hand, we cannot reject the J-test for HAC weights at a 5 percent significant level. Despite primarily interested in results corrected for heteroscedascity and autocorrelation, we have included estimation of other weights, to highlight this difference. Mavroeidis (2005) argues that HAC weights may lower the power of the J-test (see below) and we should thus not assure ourselves that the moment conditions may not be violated.

The GMM residual is given as $e_t = \eta_t - \alpha_3 (\pi_{t+1} - E(\pi_{t+1} \mid I_t))$ from (3.1.16), It consists of a structural error η_t from (3.1.13) and a forecast error $(\pi_{t+1} - E(\pi_{t+1} \mid I_t))$. Thus, a violation of the moment conditions might be caused by a failure of rationality $E((\pi_{t+1} - E(\pi_{t+1} \mid I_t)) z_t) \neq 0$ or caused by an omitted variable problem $E(\eta_t z_t) \neq 0$. Due to lack of independent data on expectations we are not able to identify which of the interpretations cause the violation of the moment condition.

Overinstrumenting and HAC estimator We note that using an autocorrelation consistent weight matrix the J-test is insignificant and the overidentifying conditions are not rejected. Nielsen (2019b, sec. 7.3) makes the same observation.

Mavroeidis (2005) finds that in finite samples the power of the J-test is significantly affected by applied HAC-weight that account for serial correlation and heteroscedasticity in residuals. Mavroeidis (2005) suggests that this might partially be caused by a slower convergence of a HAC(12) estimator to the true asymptotic variance of moment conditions than e.g. MA-l estimator. This paper has applied a HAC(12) from Newey and West (1987) as implemented in OxMetrics. It is considered outside the scope of this paper to apply other serial correlation-correcting estimators.

Mavroeidis (2005) further adds that overinstrumenting, i.e. adding irrelavant instruments, reduces the power of the J-test. He adds that this is further troubled by lack of testing for *weak* identification in forward-looking models. For this reason we've included only rather 'recent' lags compared to e.g. Galí and Gertler (1999). Verbeek (2008) and Nielsen (2019b) make no or little notice of these issues and thus any considerations here will only be qualitative.

3.2.3 Hybrid model

We wish to extend the model with a backward term for inflation:

$$\pi_t = \alpha_1 + \alpha_2 u_t + \alpha_3 \pi^e_{t+1|t} + \alpha_4 \pi_{t-1} + \eta_t, \quad t = 1, \ldots, T \tag{3.2.1}$$

In accordance to section 3.2.1 we reformulate equation (3.2.1) due to latency of $E(\pi_{t-1} \mid I_t)$:

$$\pi_t = \alpha_1 + \alpha_2 u_t + \alpha_3 \pi_{t+1} + \alpha_4 \pi_{t-1} + e_t \tag{3.2.2}$$

We instrument inflation π_{t+1} with up to four lagged values. The backward term for inflation π_{t-1} is still included in the list of instruments, and is as such an instrument for itself, likewise for the measure of output gab[5] (unemployment-rate deviation from mean) u_t. The estimation results for the hybrid-model (H) are to be shown in table 3.6.

Results for the Hybrid model

Estimation results of model (3.2.2) for the π_{t+1} and π_{t-1} are all positive and significant at a 5 pct. significant level, with levels of approximately 1/2. Estimates for the u_t are however insignificant on a 5 pct. level. Results are fairly robust to changes in estimator (iterated and two-step GMM)

We do not reject any of the J-tests for the model (3.2.2) indicating that moment conditions are not violated (prior considerations taken into account).

Does the Phillips curve seems to be primarily forward looking or backward looking?

Accounting for standard errors the estimates for α_3 and α_4 are close to equality, and we do not find evidence of either being more dominant.

We refer to Galí and Gertler (1999) for theoretical deduction of the model, however note that the sum of coefficients are close to unity, consistent with the understanding that the backward and forward term represent shares of firms with different price setting behaviours. When restricting the sum of $\alpha_3 + \alpha_4 = 1$ we find results do not deviate.

Galí and Gertler (1999) find that the backward model is dominant ($\alpha_4 > \alpha_3$), which is not reproducible in the current estimation. Differences might be due to issues with instrumenting (noted by Mavroeidis (2005)) and different data frequency and measures for output gap.

Conclusions are robust to subdividing the sample.

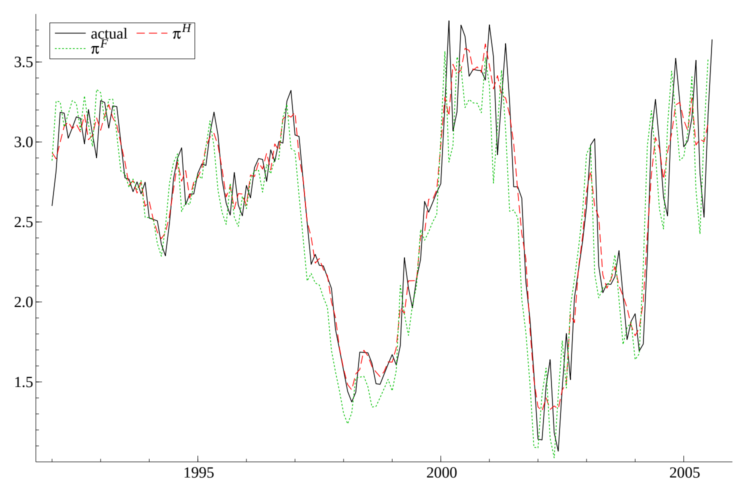
Goodness of fit Despite no coefficient being dominant of the other, we might visually illustrate that the backward-model seems to fit actual data better than the forward-looking model, which does not sufficiently account for the persistence in inflation, see figure 3.5.

However, not being able to find a real effect of output gap on inflation is unsettling for the basic story – however this might be explained by the measure of output gap. For instance using capital utilisation's deviations from mean result in real effects as well as the forward looking model being most dominant (in line with Gali et al. (2005)).

[5]Note that in section 3.1.1 we argued for the exogeneity of u_t and will comply herewith, however note that Galí and Gertler (1999) treats the measure as endogenous. We note that it does not affect results remarkably.

Figure 3.5: Actual inflation and calculated inflation

(a) Full sample (1971-2005)



(b) Subsample (1992-2005)

Notes: Inflation is calculated based on estimates from table 3.6 lines 3 and 9, using iterated GMM and HAC weights for forward-looking and hybrid model respectively. Subsample is shown for better visualisation.

Appendix

Derivation of (3.1.5)

We can rewrite the numerator so it is clear that $\hat{\alpha}_3$ only is consistent if there is no measurement error.

$$
\begin{aligned}
E(\pi^{\text{survey}}_{t+1|t}\epsilon_t) &= E[(\pi^e_{t+1|t} + v_t)(\eta_t - \alpha_3 v_t)] \\
&= E[\pi^e_{t+1|t}\eta_t - \alpha_3 v_t \pi^e_{t+1|t} + v_t\eta_t - \alpha_3 v_t^2]
\end{aligned}
$$

We use that the expectation is a linear operator:

$$
= E(\pi^e_{t+1|t}\eta_t) - \alpha_3 E(v_t \pi^e_{t+1|t}) + E(v_t\eta_t) - \alpha_3 E(v_t^2)
$$

We use that η_t and v_t are independent of relevant variables:

$$
= E(\pi^e_{t+1|t})E(\eta_t) - \alpha_3 E(v_t)E(\pi^e_{t+1|t}) + E(v_t)E(\eta_t) - \alpha_3 E(v_t^2)
$$

We apply $E(v_t) = E(\eta_t) = 0$:

$$
\begin{aligned}
&= -\alpha_3 E(v_t^2) \\
&= -\alpha_3 V(v_t)
\end{aligned}
$$

Bibliography

Comincioli, B. (1996). The stock market as a leading indicator: An application of granger causality. *University avenue undergraduate journal of economics 1*(1), 1.

David Gulley, O. and J. Sultan (1998). Consumer confidence announcements: do they matter? *Applied financial economics 8*(2), 155–166.

Doornik, J. A. and D. F. Hendry (2018a). *PcGive 15*, Volume 2. Timberlake.

Doornik, J. A. and D. F. Hendry (2018b). *PcGive 15*, Volume 1. Timberlake.

Foresti, P. (2006). Testing for Granger causality between stock prices and economic growth. *MPRA Paper* (2962), 1–10.

Gali, J., M. Gertler, and J. D. Lopez-Salido (2005). Robustness of the estimates of the hybrid new keynesian phillips curve. *Journal of Monetary Economics 52*(6), 1107–1118.

Galí, J. and M. Gertler (1999). Inflation dynamics: A structural econometric analysis. *Journal of monetary Economics 44*(2), 195–222.

Kim, S.-J. (1999). Do macro-economic news announcements affect the volatility of foreign exchange rates? some evidence from australia. *Applied Economics 31*(12), 1511–1521.

Mavroeidis, S. (2005). Identification issues in forward-looking models estimated by gmm, with an application to the phillips curve. *Journal of Money, Credit and Banking 37*(3), 421–448.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix. *Econometrica 55*(3), 703–708.

Nielsen, B. H. (2019a). Analysis of vector autoregressive models. Lecture Note 3.

Nielsen, B. H. (2019b). Generalized method of moments estimation. Lecture Note 7.

Nielsen, B. H. (2019c). Introduction to vector and matrix diffrentiation. Appendix.

Nielsen, B. H. (2019d). Modelling volatility in financial time series: An introduction to arch. Lecture Note 6.

Verbeek, M. (2008). *A guide to modern econometrics.* John Wiley & Sons.