Exam Solution Guide

Econometrics II

May 2017

Part 1
Forecasting the US Inflation Rate

The Case The goal of this part of the exam is to estimate a univariate time series model for the US inflation rate over the period from 1960:1 to 2016:4 and use the estimated model to forecast the inflation rate until 2026:4.

The Data Graphs of the data and relevant transformations must be shown in the exam. It must be noted that the level of the inflation rate is clearly non-stationary, but the first-difference appears somewhat stationary. It can be noted that the level of the variance of the change in the inflation rate appears to shift around 1970 and again around 1990.

Econometric Theory The econometric theory must include the following:

(1) A precise definition and interpretation of the model considered and its properties. Specifically, a univariate autoregressive (AR) or autoregressive moving average (ARMA) model must be presented. Furthermore, a precise definition of the out-of-sample forecasts and the forecast variance must be given.

(2) A precise description of the estimator used, in particular a precise account of the assumptions used to derive the estimator. Specifically, the method of moments (MM) or the maximum likelihood (ML) estimators can be used dependent on the model considered.

(3) A precise account of the necessary assumptions for consistent estimation and valid inference. This includes a precise definition of the null hypotheses, test statistics, and asymptotic distributions used to test relevant hypotheses.

(4) The theory must be presented precisely and in a logical order with a consistent and correct notation.

Empirical Results The empirical results must include the following:

(1) A presentation and discussion of the relevant empirical results, so that the reader is able to understand the steps carried out in the process as well as the conclusions made.

(2) A description of the model selection process based on a general-to-specific approach, information criteria, or both.

(3) A discussion of whether the assumptions for consistent estimation and valid inference are satisfied for the estimated models. Specifically, this includes misspecification testing, which must be presented and discussed before statistical testing is carried out.

(4) A clear conclusion to the main question and a discussion of the limitations of the approach used to reach the conclusion. Specifically, the conclusion regarding the out-of-sample forecasts and the forecast variance must be presented and the limitations of the estimated models must be discussed in relation the forecasts.

(5) When the forecast is presented, it should be emphasized that aim is to forecast the level of the inflation rate and not the changes in the inflation rate.

Part 2
Are oil prices cointegrated?

The Case The goal of this part of the exam is to use cointegration techniques to test the empirical validity of a theoretical equilibrium relation between two oil prices, WTI and Brent.

The Data Graphs of the data and relevant transformations must be shown in the exam. It must be noted that the time series for WTI and Brent (log-)prices appear to be unit root processes with some long-run co-movements indicating cointegration between them. One may also include a graph of the spread between the two prices. Roughly, this spread appears to be stationary.

Econometric Theory The econometric theory must include the following:

(1) A precise definition and interpretation of the models considered and their properties. Specifically, an interpretation of cointegration must be presented along with a presentation of univariate autoregressive (AR) models used to test for unit roots and a single equation cointegration approach based on the Engle-Granger two-step procedure or the autoregressive distributed lag (ADL) and error-correction models (ECM).

(2) A precise description of the estimator used, in particular a precise account of the assumptions used to derive an estimator.

(3) A precise account of the necessary assumptions for consistent estimation and valid inference. This includes a precise definition of the null hypotheses, test statistics, and asymptotic distributions used to test relevant hypotheses.

(4) The theory must be presented precisely and in a logical order with a consistent and correct notation.

Empirical Results The empirical results must include the following:

(1) A presentation and discussion of the relevant empirical results, so that the reader is able to understand the steps carried out in the process as well as the conclusions made.

In light of the fact that the spread between the two prices seems stationary, it can be noted that the cointegration analysis might be carried out using a known cointegration vector $(1, -1)'$.

(2) A description of the model selection process based on a general-to-specific approach, information criteria, or both.

(3) A discussion of whether the assumptions for consistent estimation and valid inference are satisfied for the estimated models. Specifically, this includes misspecification testing, which must be presented and discussed before statistical testing is carried out.

(4) A clear conclusion to the main question and a discussion of the limitations of the approach used to reach the conclusion. Specifically, the conclusion regarding cointegration between the two prices must be presented and the limitations of the single-equation cointegration approach must be discussed in relation to the conclusion.

Part 3
Friday effects in stock market volatility

The Case The goal of this part of the exam is to investigate if there is a so-called Friday effect in stock market returns. Analysis should be carried out using a (G)ARCH model.

The Data Graphs of the log-returns and relevant transformations must be shown in the exam. It must be noted that the magnitude of returns varies over time, indicating that the returns are conditionally heteroskedastic.

Econometric Theory The econometric theory must include the following:

(1) A precise definition and interpretation of the model considered and its properties.

(2) A precise description of the test for no ARCH.

(3) A precise description of the estimator used, in particular a precise account of the assumptions used to derive an estimator.

(4) A precise account of the necessary assumptions for consistent estimation and valid inference. This includes a precise definition of the null hypotheses, test statistics, and asymptotic distributions used to test relevant hypotheses.

(5) The theory must be presented precisely and in a logical order with a consistent and correct notation.

Empirical Results The empirical results must include the following:

(1) A presentation and discussion of the relevant empirical results, so that the reader is able to understand the steps carried out in the process as well as the conclusions made.

(2) A discussion of whether the assumptions for consistent estimation and valid inference are satisfied for the estimated models.

(3) A precise discussion/judgement of potential misspecification of the model.

(4) A conclusion to the main question. Specifically, if there is empirical evidence of the Friday effect. As a minimum, it should be investigated if there is a Friday effect in the conditional volatility. Some might choose to include the Friday dummy in the equation for the conditional mean of the returns.

(5) A clear conclusion to the main question and a discussion of the limitations of the approach used to reach the conclusion.

Part 4
Theoretical Problem:
Estimation of ARMA Models

Estimation of ARMA models

Consider the ARMA(1,1) model

$$\begin{aligned} y_t &= \phi y_{t-1} + u_t, \quad t = 1, ..., T \\ u_t &= \varepsilon_t + \alpha \varepsilon_{t-1}, \end{aligned}$$

where $\varepsilon_t \sim IIDN(0, \sigma^2)$, and given some initial values y_0 and $\varepsilon_0 = 0$.
Given a set of observations $(y_1, ..., y_T)$, we seek to estimate the model parameters $\theta = (\phi, \alpha, \sigma^2)'$.
Let $\theta_0 = (\phi_0, \alpha_0, \sigma_0^2)'$ denote the vector of true parameters that have generated the data.
Unless stated otherwise, we assume throughout that $|\phi_0| < 1$, $|\alpha_0| < 1$, and $\phi_0 \neq -\alpha_0$.

Question 1
In order for the OLS estimator to be consistent, we need that $E[u_t y_{t-1}] = 0$. This condition is not satisfied if $\alpha_0 \neq 0$. Hence, in general, the OLS estimator for ϕ is inconsistent.
Detailed arguments should be included. Ideally, a formal proof for the inconsistency is provided. This relies on a law of large numbers, which applies here, since the data generating process is stationary, as $|\phi_0| < 1$.

Question 2
As ε_t is assumed to be normally distributed, an estimate of θ can be obtained by maximum likelihood estimation, where the likelihood function is based on the density of the normal distribution. The arguments are very similar to the ones used for deriving the likelihood function for the MA(1) model.

Conditional on y_0 and $\varepsilon_0 = 0$, we have that the joint density of $(y_1, ..., y_T)$ is

$$f(y_1, ..., y_T | y_0, \varepsilon_0 = 0) \quad = \quad \prod_{t=1}^{T} f(y_t | y_{t-1}, ..., y_1, y_0, \epsilon_0 = 0),$$

where

$$f(y_t|y_{t-1},...,y_1,y_0,\epsilon_0=0) \quad = \quad \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{\epsilon_t^2}{2\sigma^2}\right\},$$

with ε_t given recursively as

$$\begin{aligned}
\varepsilon_0 &= 0 \\
\varepsilon_1 &= y_1 - \phi y_0 - \alpha\varepsilon_0 \\
&\vdots \\
\varepsilon_t &= y_t - \phi y_{t-1} - \alpha\varepsilon_{t-1} \\
&\vdots \\
\varepsilon_T &= y_T - \phi y_{T-1} - \alpha\varepsilon_{T-1}.
\end{aligned}$$

Based on this (conditional) joint density, we obtain the likelihood function.
Detailed arguments should be provided. In particular, the (log-)likelihood function should be stated.
Moreover, it should be mentioned that we cannot find a closed-form expression for the maximum likelihood estimator. The estimator is obtained by numerical methods.

Question 3
We have the MA-type representation

$$y_t \quad = \quad \sum_{i=0}^{t-1}\phi_0^i u_{t-i} + \phi_0^t y_0.$$

Since ε_t is an IID process and $u_t = \varepsilon_t + \alpha_0\varepsilon_{t-1}$, it holds that u_t and u_{t-j} are independent for $j \geq 2$. Hence $E[u_t u_{t-j}] = 0$ for $j \geq 2$. In light of the MA-type representation, it follows that $E[u_t y_{t-j}] = 0$ for $j \geq 2$.

Based on the moment condition, we can use y_{t-2} as an instrument. An IV estimator for ϕ is

$$\widehat{\phi}_{IV} \quad = \quad \frac{\sum_{t=2}^T y_{t-2}y_t}{\sum_{t=2}^T y_{t-1}y_{t-2}}.$$

Detailed derivations should be provided.

Note that the IV estimation relies on having that y_{t-2} is valid and relevant. The validity has already been established. The instrument is not relevant if $E[y_{t-1}y_{t-2}] = 0$, but this is ruled out by the assumption that $\phi_0 \neq -\alpha_0$.

Alternatively, one may estimate ϕ based on GMM where several instruments are used, e.g. $z_t = (y_{t-2},...,y_{t-q})'$ for some $q \geq 3$.

Question 4

One may start out by noting, that the ML estimator is expected to have the smallest variance. This is the case if

$$\frac{\Omega_{IV}}{\Omega_{ML}} \geq 1.$$

It holds that

$$\begin{aligned}\frac{\Omega_{IV}}{\Omega_{ML}} &= \frac{(1+4\alpha_0^4+4\phi_0\alpha_0+4\phi_0\alpha_0^3+2\alpha_0^2\phi_0^2+\alpha_0^2)}{(1+\phi_0\alpha_0)^4} \\ &= \frac{(1+4\alpha_0^4+4\phi_0\alpha_0+4\phi_0\alpha_0^3+2\alpha_0^2\phi_0^2+\alpha_0^2)}{(1+\phi_0^4\alpha_0^4+4\alpha_0^2\phi_0^2+2\alpha_0^2\phi_0^2+4\alpha_0\phi_0+4\phi_0^3\alpha_0^3)}.\end{aligned}$$

In order to have $\frac{\Omega_{IV}}{\Omega_{ML}} \geq 1$, the difference between the numerator and the denominator has to be non-negative. The difference (i.e. the numerator minus the denominator) is equal to

$$4\alpha_0^4+4\phi_0\alpha_0^3+\alpha_0^2-\phi_0^4\alpha_0^4-4\alpha_0^2\phi_0^2-4\phi_0^3\alpha_0^3 \;=\; \alpha_0^2(1-\phi_0^4\alpha_0^2)+4\phi_0\alpha_0^3(1-\phi_0^2)+4\alpha_0^2(\alpha_0^2-\phi_0^2).$$

Note that all the terms on the right-hand side are strictly positive, under the condition $0 < \phi_0 \leq \alpha_0 < 1$. So, under this condition, we have that the asymptotic variance of the ML estimator is smaller than the one of the IV estimator, and we may conclude that the ML estimator is the best estimator.

Question 5

In the case $\phi_0 = -\alpha_0$, both variances are infinite. Loosely speaking, we may interpret this as the parameter cannot be estimated with any precision.

The reason for this is that the parameter is not identified. To see this, with L the lag operator, define $\phi(L) = 1 - \phi_0 L$ and $\alpha(L) = 1 + \alpha_0 L$. The ARMA(1,1) process is then

$$\phi(L)y_t \;=\; \alpha(L)\varepsilon_t.$$

Note that if $\phi_0 = -\alpha_0$, then $\phi(L) = \alpha(L)$, and provided that the lag polynomials are invertible, for *any* values of $\phi_0 = -\alpha_0$, $y_t = \varepsilon_t$. Hence the process is a white noise process, even if the true parameters are different from zero.