Final Exam Financial Econometrics

LAST, first name: ____

June 2016

Hereby I confirm that during the exam I complied with the EPFL honor code

Signature:______________________________

- Write as clearly as possible.
- Non-programmable calculators are allowed.
- The exam is closed book, closed notes.
- Hand in the exam in 3h.
- The weighting of exercises is indicative and may be subject to changes.
- The final exam is organized as follows:
  - Exercise 1: Theory (pp. 2-3);
  - Exercise 2: Interpretation (pp. 4-8).
  - Exercise 3: Questions related to the course (pp. 9-10);
- **PLEASE HAND IN THE TEXT OF THE EXAM WITH YOUR COPY**.

Exercise 1: VAR(1) process ($\simeq 40\%$)

Consider the following VAR(1) representation:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 1 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}.$$

Using matrix notation, it can be written as:

$$X_t = AX_{t-1} + \varepsilon_t$$

where $X_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix}$ and $\varepsilon_t = \begin{pmatrix} u_t \\ v_t \end{pmatrix}$. We further assume that $(\varepsilon_t, t \in \mathbb{Z})$ is a weak white noise (vector) with variance-covariance matrix $\Sigma = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$. In the sequel, if $(z_t, t \in \mathbb{Z})$ is some univariate or vector-valued time series, then $\underline{z_t}$ and $\mathbb{EL}\left[. \mid \underline{z_t}\right]$, respectively, denote the information available at time t, $\underline{z_t} = \{z_s, s \leq t\}$, and the best linear predictor given the information set $\underline{z_t}$.

Part I: Stationary solution and innovation process

1. Check that the model has a stationary solution.
2. Show that $(\varepsilon_t, t \in \mathbb{Z})$ is the innovation process of $(X_t, t \in \mathbb{Z})$.

Part II: Specification of $(y_t, t \in \mathbb{Z})$ as an `ARMA` process

3. Using the VAR specification, show that $(y_t, t \in \mathbb{Z})$ can be written as:

$$\Phi(L)y_t = w_t$$

   where $\Phi(L) = 1 - \frac{7}{6}\mathtt{L} + \frac{1}{3}\mathtt{L}^2$ and w_t is a linear function of v_t, v_{t-1} and u_{t-1}.

4. Using the previous result (Question 3),
   - Find the variance of w_t.
   - Show that the autocovariance function of $(w_t, t \in \mathbb{Z})$ satisfies $\gamma_w(h) = 0$ for all $|h| \geq 1$.

   In the sequel, we assume that (w_t) is a weak white noise.

5. Using the infinite order moving average representation (`MA`(∞)) and the results of Part II, determine the sequence of real numbers (c_k) such that:

$$y_t = \sum_{k=0}^{+\infty} c_k w_{t-k}.$$

6. Using only the `MA`(∞) representation (Question 5), find the autocovariances $\gamma_y(0)$, $\gamma_y(1)$, and $\gamma_y(h)$ (for a given $h > 0$) in terms of the c_k coefficients without explicitly considering the numerical values.

7. Using the results of Question 3, find the best linear predictor of y_{T+1} and y_{T+2} given the information set $\underline{y_T}$, denoted $\hat{y}_T^\star(1)$ and $\hat{y}_T^\star(2)$, and the corresponding forecasting errors, $e_{y,T}(1)$ and $e_{y,T}(2)$.

8. Find the difference equation that defines the best linear predictor of y_{T+h} (for $h \geq 3$) given the information available at time T. Write down the general solution (without determining the constants of this general solution). What happens as $h \to \infty$?

Part III: Forecasts and information sets for $(y_t, t \in \mathbb{Z})$

9. Find $\mathbb{EL}\left[y_t \mid \underline{x_{t-1}}, \underline{y_{t-1}}\right]$.

10. Show that:

$$\mathbb{V}\left[y_t - \mathbb{EL}\left[y_t \mid \underline{x_{t-1}}, \underline{y_{t-1}}\right]\right] < \mathbb{V}\left[y_t - \mathbb{EL}\left[y_t \mid \underline{y_{t-1}}\right]\right].$$

11. Explain the intuition of the previous result (Question 10).

Part IV: Forecasts and information sets for $(x_t, t \in \mathbb{Z})$

12. Find the best linear approximation of u_t, denoted $u_t^\star$, defined by $u_t^\star = \mathbb{EL}\left[u_t \mid v_t\right]$.

13. Using the previous result (Question 12), find $\mathbb{EL}\left[x_t \mid \underline{x_{t-1}}, \underline{y_t}\right]$ and $\mathbb{EL}\left[x_t \mid \underline{x_{t-1}}, \underline{y_{t-1}}\right]$.
    Hint: Note that $\mathbb{EL}\left[x_t \mid \underline{x_{t-1}}, \underline{y_t}\right] = \mathbb{EL}\left[x_t \mid \underline{x_{t-1}}, \underline{y_{t-1}}, y_t\right]$.

14. Find $\mathbb{V}\left[x_t - \mathbb{EL}\left[x_t \mid \underline{x_{t-1}}, \underline{y_{t-1}}\right]\right]$ and $\mathbb{V}\left[x_t - \mathbb{EL}\left[x_t \mid \underline{x_{t-1}}, \underline{y_t}\right]\right]$, and compare them.

Exercise 2: Interpretation ($\simeq 30\%$)

Question 1

Suppose that you are estimating the following time series regression model (Model 1) for the period January 2001 to November 2015 (178 monthly observations):

$$z_t = \alpha + \beta z_{m,t} + u_t$$

where z_t and $z_{m,t}$ denote the excess (log-)return (in percent) of a given stock and the excess market (log-)return (in percent), respectively. Table 1 provides the ordinary least squares estimates of Model 1.

Table 1: OLS estimation of Model 1

Variable	Estimate	Std. Error	P-value
Constant	15.015	1.7677	0.0000
z_m	1.491	0.0836	0.0000
R^2	0.6433	$\bar{R}^2$	0.6413
s_u	23.353		

Note: s_u is the standard deviation of the residuals.

1.1. Comment on the estimation results, and especially on the magnitude, statistical significance and interpretation of the estimates.

1.2. What is the proportion of the risk that can be diversified? Explain carefully.

1.3. The average monthly annualized return of the studied stock over the estimation period is 14.24% with a standard deviation of 38.66%. Over the same period, the average yield on a 10-year government bond (the riskless asset) was 3.56%. Using these figures, as well as the results reported in Table 1, calculate and provide an interpretation of the corresponding Sharpe ratio (i.e., the mean excess return to volatility ratio), Treynor ratio (i.e. the mean excess return to beta ratio) and Jensen's alpha.

1.4. The model is re-estimated using the following specification (Model 2):

$$\begin{aligned} z_t &= \alpha + \beta z_{m,t} + u_t \\ u_t &= \sigma_t w_t \\ \sigma_t^2 &= \omega + \gamma_1 u_{t-1}^2 + \phi_1 \sigma_{t-1}^2 \end{aligned}$$

where $w_t \sim \texttt{i.i.d.}\mathcal{N}(0,1)$. Table 2 displays the estimation results (using maximum likelihood estimation).

Table 2: Maximum likelihood estimation of Model 2

Parameter	Estimate	Std. Error	P-value
α	6.665	0.943	0.000
β	1.871	0.049	0.000
ω	30.791	12.914	0.017
γ_1	0.888	0.227	0.000
ϕ_1	0.108	0.088	0.000

- Explain carefully the features of financial returns series the specification of the error terms aims to capture.
- Comment on the estimation results, and especially on the magnitude, statistical significance and interpretation of the estimates.
- Determine the unconditional variance and interpret it.

Question 2

One would like to capture the dynamics of the French 10-year government bond yield (`GBY`) using an `ARIMA` specification.

2.1. In this respect, two Augmented Dickey-Fuller tests have been conducted on the French 10-year government bond yield (`GBY`) and its first difference (Δ`GBY`) for the period January 2000 - October 2015 (187 monthly observations). In both cases, one makes use of a first difference transformation of the dependent variable. In addition, the maximum number of lags, which is determined by some information criteria, is equal to four in both cases. Finally, only a constant term is included in both specifications. Table 3 reports the value of the Student-based test statistic for the two tests as well as the critical values at 1%, 5%, and 10% level.

Table 3: ADF tests

`GBY`	
Test statistic	-0.4244
Δ`GBY`	
Test statistic	-9.6751
Critical values	
1% level	-3.4653
5% level	-2.8768
10% level	-2.5750

- Write down the test regression in both cases (i.e., for `GBY`).
- Write down the null and alternative hypothesis.
- What is your conclusion? Explain carefully.

2.2. Three models are estimated for ΔGBY over the period February 2000 - October 2015. Table 4 provides the (maximum likelihood) estimation results and the values of the information criteria.

Table 4: Selection of univariate time series models

Parameter	Estimate	Std. Error	P-value	Inf. Criterion	Value
			$\texttt{AR}(2)$		
Constant	-0.0243	0.0135	0.0719	AIC	-0.847
				SBIC	-0.795
ϕ_1	0.3032	0.0727	0.0000	HQ	-0.826
ϕ_2	-0.1572	0.0727	0.0318		
			$\texttt{MA}(1)$		
Constant	-0.0255	0.0153	0.0974	AIC	-0.862
				SBIC	-0.828
θ_1	0.3486	0.0687	0.0000	HQ	-0.848
			$\texttt{ARMA}(1,2)$		
Constant	-0.0320	0.0066	0.0000	AIC	-0.956
				SBIC	-0.887
ϕ_1	0.9087	0.0354	0.0000	HQ	-0.928
θ_1	-0.6512	0.0753	0.0000		
θ_2	-0.4271	0.0740	0.0000		

Notes: (1) ϕ_1 (resp., ϕ_2) is the first (respectively, second) autoregressive coefficient. (2) θ_1 (respectively, θ_2) is the moving average coefficient of the first (respectively, second) lag of the error term. (3) AIC, SBIC and HQ stand for the Akaike, Schwarz-Bayesian, and Hannan-Quinn information criteria.

- Write down each model and explain the intuition underlying their specification (i.e., the use of an autoregressive part and/or a moving average part).
- Find the unconditional mean of each model.
- Explain carefully which model is the most suitable to model ΔGBY.

Question 3

One would like to study the relationship between the USD/GBP log-spot exchange rate, denoted `SPOT_ER`, and the (twelve-month) USD/GBP log-forward exchange rate, denoted `FORWARD_ER`, over the sample period January 2001 - October 2015.

3.1. In a first step, unit root tests provide evidence that the two variables are integrated of order one. In a second step, an OLS estimation of the following time series regression is considered:

$$\texttt{SPOT\_ER}_t = \alpha + \texttt{FORWARD\_ER}_{t-12}\beta + u_t.$$

Table 5 reports the estimation results.

Table 5: OLS estimation

Variable	Estimate	Std. Error	P-value
Constant	0.0984	0.0137	0.000
FORWARD_ER	0.5604	0.0622	0.000
R^2	0.3154	$\bar{R}^2$	0.3115

In addition, an Augmented Dickey-Fuller test is conducted on the residuals. The student-based test statistic is -3.7834—the critical values being -3.468 (1%), -2.878 (5%), and -2.575 (10%).

- What does this relationship capture? Explain carefully.
- Interpret the result of the ADF test.

3.2. In a third step, the following time series regression is estimated:

$$\begin{aligned}\Delta\texttt{SPOT\_ER}_t &= \delta_0 + \Delta\texttt{SPOT\_ER}_{t-1}\delta_{1,1} + \Delta\texttt{SPOT\_ER}_{t-14}\delta_{1,2} + \Delta\texttt{FORWARD\_ER}_{t-1}\delta_{2,1} + \Delta\texttt{FORWARD\_ER}_{t-2}\delta_{2,2} \\ &\quad +\lambda\hat{u}_{t-1} + v_t\end{aligned}$$

where $\hat{u}_{t-1}$ is the residual at time $t-1$ using the second step (Question 3.1). Table 6 reports the OLS estimation results.

Table 6: Maximum likelihood estimation of Model 2

Variable	Estimate	Std. Error	P-value
Constant	9×10^{-5}	0.0005	0.8604
$\Delta\texttt{SPOT\_ER}_{t-1}$	-0.4581	0.1192	0.0002
$\Delta\texttt{SPOT\_ER}_{t-14}$	0.1346	0.0577	0.0210
$\Delta\texttt{FORWARD\_ER}_{t-1}$	0.7906	0.0834	0.0000
$\Delta\texttt{FORWARD\_ER}_{t-2}$	0.2935	0.0785	0.0003
$\hat{u}_{t-1}$	-0.0334	0.0156	0.0343
R^2	0.4672	$\bar{R}^2$	0.4477

- What does this model aim to capture? Explain carefully.
- Comment on the estimation results, and especially on the interpretation of the coefficient of the term $\hat{u}_{t-1}$, and the interpretation and statistical significance of the estimates of $\delta_{1,1}$ and $\delta_{2,1}$.

Exercise 3: Questions related to the course ($\simeq 30\%$)

Precise and concise answers are expected.

Question 1

One would like to regress a financial variable on a very large set of explanatory variables. As a result, a problem of multicolinearity is encountered. A student suggests that one possible solution is to implement a data reduction technique, and especially a principal component analysis using the variance-covariance matrix.

- Explain how one can get the first two principal components. Precise carefully all notation.
- What is the interpretation of the eigenvalues?

Question 2

What are the three forms of the market (informational) efficiency hypothesis? Explain carefully.

Question 3

One would like to study the short-term and long-term predictive content of the dividend to price ratio for future (cumulative) (log-)returns. Taking that the dividend to price ratio is a persistent variable, one argues that "the short-term and long-term regressions are two sides of the same coin!".

- Could you explain this sentence?
- Propose a statistical model that might explain this feature of predictive regressions.

Question 4

Consider the following constant (log-)return model with `EGARCH`(1,1) error terms:

$$\begin{aligned} r_t &= \mu + \epsilon_t \\ \ln(\sigma_t^2) &= \omega + \beta_1 \ln(\sigma_{t-1}^2) + \alpha_1 \left| \frac{\epsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma_1 \frac{\epsilon_{t-1}}{\sigma_{t-1}}. \end{aligned}$$

Table 7 reports the estimation results.

Table 7: Estimation of an EGARCH(1,1)

Parameter	Estimate	Standard error	P-value
μ	0.0268	0.0091	0.0000
ω	-0.0918	0.0038	0.0000
α_1	0.1205	0.0048	0.0000
γ_1	-0.0791	0.0034	0.0000
β_1	0.9850	0.0010	0.0000

- What are the main features of this model?
  Hint: Compare with a constant (log-)return model with GARCH(1,1) error terms.
- What is the effect of a negative shock? Explain carefully.
- Using Table 8, propose an estimate of this effect.

Question 5

Table 8 reports the results of the Jarque-Berra tests for the daily (respectively, weekly) residuals of the constant (log-)return model (Model 1) using the S&P500 (US) index and the FT all shares (UK) index, the standardized residuals of a constant (log-) return model ($r_t = \mu + \epsilon_t$ for all t) with GARCH(1,1) error terms (Model 2), and the standardized residuals of a constant (log-) return model with GJR error terms (Model 3).

Table 8: Normality tests

	$S\&P500$		FT All Shares	
	Statistic	P-value	Statistic	P-value
		Daily returns		
residuals of Model 1	26861.511	0	20438.469	0
GARCH standardized residuals	3190.347	0	11641.998	0
GJR standardized residuals	3138.314	0	7121.563	0
		Weekly returns		
Residuals of Model 1	1007.024	0	1100.524	0
GARCH standardized residuals	274.235	0	373.379	0
GJR standardized residuals	135.497	0	256.735	0

Note: The GJR model specifies the conditional variance as $\sigma_t^2 = \omega + \left[\alpha_1 \epsilon_{t-1}^2 + \gamma_1 \Pi_{t-1}^- \epsilon_{t-1}^2\right] + \beta_1 \sigma_{t-1}^2$ where Π_t^- equals one if $\epsilon_t < 0$, and 0 otherwise.

- Comment on Table 9.
  Hint: Write down the null and alternative hypothesis, and then answer the following questions: (1) Can one reject the null hypothesis of normality irrespective of the frequency?; (2) Why does one observe decreasing test statistics for the log-returns (respectively, the $GARCH(1,1)$ and GJR standardized residuals)?; (3) Which salient feature of a GARCH(1,1) model might explain that one would expect the value of the test-statistic to drop relative to the one of the constant return model ?; (4) How could you explain the differences between the GARCH- and GJR-based test?

PLEASE HAND IN THE TEXT OF THE EXAM WITH YOUR COPY