Empirical Finance: Study Questions Week 1
Imperial College London
Hormoz Ramian

Question 1

(1.1) "Failure to reject H_0 means the null hypothesis is true", true or false? If true, explain why? If false, explain why.

(1.2) Is the statement, "A matrix is a projection matrix iff it is an idempotent matrix", true? If so, explain why? If not, explain when this can be true.

(1.3) "An idempotent matrix is always invertible", true or false?

(1.4) "A projection matrix is always invertible", true or false?

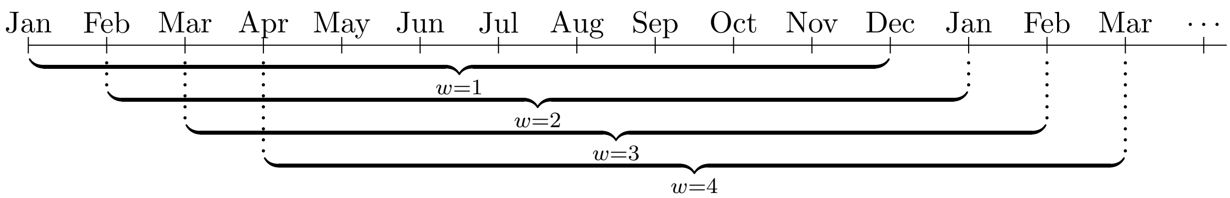
(1.5) In March 1994, Michigan voters approved a sales tax increase from 4% to 6%. In political advertisements, supporters of the measure referred to this as a two percentage point increase, or an increase of two cents on the dollar. Opponents of the tax increase called it a 50% increase in the sales tax rate. Explain which way of measuring the increase in the sales tax is more accurate.

Question 2 Suppose the sample size is T. We wish to estimate parameters $\{\beta_w\}_{w=1}^{T-N+1}$,

$$y_t = x_t\beta_w + u_t$$

within smaller subsamples of size N with $N < T$ such as monthly returns from January to December

- denote each regression with w
- there exists $T-N+1$ (overlapping) subsamples
- first rolling regression ($w = 1$) uses observations e.g. from 1 to 12 (January to December)
- second ($w = 2$) uses 2 to 13 (February to January next year), and so on
- till the end of the sample $(T - 11)$ to T
- each regression gives a $\widehat{\beta}_{1,w}$



Now consider an example from the capital asset pricing model (CAPM) with the following specification used to interrelate the excess return, on a given asset $r_t - r_{f,t}$ where $r_{f,t}$ is the risk-free rate, to the market return denoted by $r_{m,t}$

$$\underbrace{r_t - r_{f,t}}_{\text{risk premium}} = \alpha_w + \beta_w(\underbrace{r_{M,t} - r_{f,t}}_{\text{market premium}}) + u_t \tag{1}$$

note that the object of interest is the rolling feature of the coefficient $\widehat{\beta}_w$. Particularly, $\widehat{\beta}_w$ summarizes the conditional relationship (given the rolling window) between the market risk premium $r_{M,t} - r_{f,t}$ and the excess return.

- expected excess returns' sensitivity to the expected excess market returns ($\widehat{\beta}_w$)
- when market risk premium increases, how much the asset return should be?

- zero value classifies investments without (low) risk, e.g. money market funds with a constant share value of \$1, certificates of deposit backed by federal deposit insurance, and cash (inflation erodes the purchasing power of money, meaning your zero-beta investment actually loses if it pays interest at less than the rate of inflation.)
- negative beta correlation means an investment moves in the opposite direction from the stock market. Can you give examples?

Use the dataset provided on IBM prices, S&P500 index, Treasury bill prices and CPI to answer the following:

(2.1) Construct real excess returns for IBM and the market index excess return.

(2.2) Write a code in R to estimate rolling parameters $\{\beta_w\}_{w=1}^{T-N+1}$.

(2.3) Write a code in R to estimate the adjusted R^2 per each window. Plot this over time and interpret the economic implication of this figure.

Question 3 Categorical variables $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$, and $\mathbf{1}_n$ each is a vector of size $n \times 1$, and that $\boldsymbol{d}_2 = \mathbf{1}_n - \boldsymbol{d}_1$ with $n = n_1 + n_2$ (n_1: number of right-handed and n_2: number of left-handed) such that:

$$d_{1,i} = \begin{cases} 1 & \text{if right-handed} \\ 0 & \text{if left-handed} \end{cases}$$

suppose:

$$\boldsymbol{y} = \boldsymbol{d}_1\gamma_1 + \boldsymbol{d}_2\gamma_2 + \boldsymbol{u}$$

(3.1) Show that $(\widehat{\gamma}_1,\ \widehat{\gamma}_2)' = (\overline{y}_1,\ \overline{y}_2)'$.

(3.2) Compare $\widehat{\gamma}_1$ and $\widetilde{\gamma}_1$ from two OLS regressions and comment:

$$\widehat{\boldsymbol{y}} = \boldsymbol{d}_1\widehat{\gamma}_1 + \boldsymbol{d}_2\widehat{\gamma}_2 \tag{2}$$

$$\widehat{\boldsymbol{y}} = \boldsymbol{d}_1\widetilde{\gamma}_1 \tag{3}$$

(3.3) Explain if the following regression can be estimated:

$$\widehat{\boldsymbol{y}} = \alpha + \boldsymbol{d}_1\widehat{\gamma}_1 + \boldsymbol{d}_2\widehat{\gamma}_2 \tag{4}$$

Question 4 The model is:

$$y_i = \beta_1 + \beta_2 x_{1i} + \beta_3 x_{2i} + u_i \tag{5}$$

At each step, state any additional assumption you need to use:

(4.1) Derive the OLS estimators without using vectors/matrix notations.

(4.2) Show that OLS estimator is unbiased.

Solutions

Question 1

(1.1) No, hypothesis testing and its conclusion are only concerned with the alternative hypothesis which is the interest of statistical inference. In general, failure to reject the null does not provide evidence that the null is actually statistically true.

(1.2) No, in general, a matrix is a projection matrix if it is an idempotent matrix but not the reverse, as projection property is only a special case of idempotency. If an idempotent matrix is symmetric, then it is also a projection matrix. A projection matrix is also known as an orthogonal projection matrix. In econometrics, we mostly work with orthogonal projection matrices and the term *orthogonal* is often dropped. However, in a more general setting we also have a definition for non-orthogonal projection matrices. Both orthogonal and non-orthogonal projection matrices are idempotent, but an orthogonal projection matrix is also necessarily symmetric[1]:

$$\boldsymbol{P} = \boldsymbol{P}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{P}' = \boldsymbol{P}'\boldsymbol{P}$$

(1.3) No, invertibility and idempotency are two separate properties. A matrix can be idempotent but not full rank.

(1.4) No, invertibility and projection properties need not coincide. A matrix can be a projection matrix but not full rank.

(1.5) Both statements are equivalent in terms of measurements. The question wishes to establish that 'percentage change' and 'percentage points change' are different ways of measuring changes (in this part, the text measures changes correctly using both methods). Naturally, each group reported the measure that made its position sound most favorable to non-econometricians.

Question 2 See the code attached.

Question 3

(3.1) In order to show that $(\widehat{\gamma}_1,\ \widehat{\gamma}_2)' = (\overline{y}_1,\ \overline{y}_2)'$, let's re-write the regression model in matrix form and compute the inverse of the quadratic form for explanatory variables. Next, note that because $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are binary variables and are perfectly orthogonal, then their inner product is zero, and the inverse form simplifies to a inverse of diagonal elements inside the matrix:

$$\begin{aligned}
\begin{pmatrix} \widehat{\gamma}_1 \\ \widehat{\gamma}_2 \end{pmatrix} &= \begin{pmatrix} \boldsymbol{d}_1'\boldsymbol{d}_1 & \boldsymbol{d}_1'\boldsymbol{d}_2 \\ \boldsymbol{d}_2'\boldsymbol{d}_1 & \boldsymbol{d}_2'\boldsymbol{d}_2 \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{d}_1' \\ \boldsymbol{d}_2' \end{pmatrix} \boldsymbol{y} \\
&= \begin{pmatrix} \sum_{i\in n_1} d_{1,i}^2 & 0 \\ 0 & \sum_{i\in n_2} d_{2,i}^2 \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{d}_1' \\ \boldsymbol{d}_2' \end{pmatrix} \boldsymbol{y} \\
&= \begin{pmatrix} n_1 & 0 \\ 0 & n_2 \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{d}_1'\boldsymbol{y} \\ \boldsymbol{d}_2'\boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \frac{1}{n_1} & 0 \\ 0 & \frac{1}{n_2} \end{pmatrix} \begin{pmatrix} \sum_{i\in n_1} y_i \\ \sum_{i\in n_2} y_i \end{pmatrix} = \begin{pmatrix} \frac{1}{n_1}\sum_{i\in n_1} y_i \\ \frac{1}{n_2}\sum_{i\in n_2} y_i \end{pmatrix} = \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix}
\end{aligned}$$

[1] A non-orthogonal projection matrix is not necessarily symmetric but not often used in statistics and econometrics.

(3.2) Under the definition of the model, we have $\widehat{\gamma}_1 = \widetilde{\gamma}_1$ since $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are perfectly orthogonal, indicating that we can split the regression for each sample and run individual regressions. Alternatively, we note that because of orthogonality, each original coefficient $\widehat{\gamma}_i$ only measures the impact from its respective variable. Algebraically, running the regression without either of the variables, amounts to no estimation bias due to omitted variable (as a result of being mutually orthogonal).

$$\widetilde{\gamma}_1 \;=\; (\boldsymbol{d}_1'\boldsymbol{d}_1)^{-1}\boldsymbol{d}_1'\boldsymbol{y} \;=\; \frac{\sum_{i\in n_1} y_i}{n_1} \;=\; \widehat{\gamma}_1 \tag{6}$$

(3.3) Let $\boldsymbol{D}_i = (1, d_{1,i}, d_{2,i})$, $\forall_i$ then $rank(\boldsymbol{D}) < 3$ because $\boldsymbol{d}_1 = \boldsymbol{1} - \boldsymbol{d}_2$ therefore $\boldsymbol{D}'\boldsymbol{D}$ is not invertible. This is because three columns of $\boldsymbol{D}$ are in perfect collinearity but if we re-write the right-hand-side with only one of the dummy variables, then the regression parameters are estimable.

Question 4

$$y_i \;=\; \beta_1 + \beta_2 x_{1i} + \beta_3 x_{2i} + u_i$$

The objective function is $RSS = \sum_{i\in N} u_i^2$ and the first order conditions (foc) are:

$$\begin{aligned}
\partial RSS/\partial\beta_1\big|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} &= -2\sum_{i\in N}\left(y_i - \widehat{\beta}_1 + \widehat{\beta}_2 x_{1i} + \widehat{\beta}_3 x_{2i}\right)\\
\partial RSS/\partial\beta_2\big|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} &= -2\sum_{i\in N} x_{1i}\left(y_i - \widehat{\beta}_1 + \widehat{\beta}_2 x_{1i} + \widehat{\beta}_3 x_{2i}\right)\\
\partial RSS/\partial\beta_3\big|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} &= -2\sum_{i\in N} x_{2i}\left(y_i - \widehat{\beta}_1 + \widehat{\beta}_2 x_{1i} + \widehat{\beta}_3 x_{2i}\right)
\end{aligned}$$

where $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3)'$. Setting the foc's equal to zero determines the following linear system which determines solution to the OLS problem:

$$0 \;=\; -2\sum_{i\in N}\left(y_i - \widehat{\beta}_1 + \widehat{\beta}_2 x_{1i} + \widehat{\beta}_3 x_{2i}\right) \tag{7}$$

$$0 \;=\; -2\sum_{i\in N} x_{1i}\left(y_i - \widehat{\beta}_1 + \widehat{\beta}_2 x_{1i} + \widehat{\beta}_3 x_{2i}\right) \tag{8}$$

$$0 \;=\; -2\sum_{i\in N} x_{2i}\left(y_i - \widehat{\beta}_1 + \widehat{\beta}_2 x_{1i} + \widehat{\beta}_3 x_{2i}\right) \tag{9}$$

The first equations gives:

$$\begin{aligned}
\sum_{i\in N} y_i &= \sum_{i\in N}\widehat{\beta}_1 + \sum_{i\in N}\widehat{\beta}_2 x_{1i} + \sum_{i\in N}\widehat{\beta}_3 x_{2i}\\
&= N\widehat{\beta}_1 + \widehat{\beta}_2\sum_{i\in N} x_{1i} + \widehat{\beta}_3\sum_{i\in N} x_{2i}
\end{aligned}$$

Dividing throughout by N and defining $\overline{x}_1 = \sum_{i\in N} x_{1i}/N$ and $\overline{x}_2 = \sum_{i\in N} x_{2i}/N$ yields:

$$\widehat{\beta}_1 = \overline{Y} - \widehat{\beta}_2\overline{x}_1 - \widehat{\beta}_3\overline{x}_2$$

which together with equations (8) and (9) gives:

$$0 \;=\; \sum_{i\in N} x_{1i}y_i - \sum_{i\in N} x_{1i}(\overline{Y} - \widehat{\beta}_2\overline{x}_1 - \widehat{\beta}_3\overline{x}_2) + \widehat{\beta}_2\sum_{i\in N} x_{1i}^2 + \widehat{\beta}_3\sum_{i\in N} x_{1i}x_{2i} \tag{10}$$

$$0 \;=\; \sum_{i\in N} x_{2i}y_i - \sum_{i\in N} x_{2i}(\overline{Y} - \widehat{\beta}_2\overline{x}_1 - \widehat{\beta}_3\overline{x}_2) + \widehat{\beta}_2\sum_{i\in N} x_{2i}x_{1i} + \widehat{\beta}_3\sum_{i\in N} x_{2i}^2 \tag{11}$$

Re-arranging gives:

$$\widehat{\beta}_2 = \frac{\operatorname{cov}(x_1, y)\operatorname{var}(x_2) - \operatorname{cov}(x_2, y)\operatorname{cov}(x_1, x_2)}{\operatorname{var}(x_1)\operatorname{var}(x_2) - [\operatorname{cov}(x_2, x_1)]^2}$$

$$\widehat{\beta}_3 = \frac{\operatorname{cov}(x_2, y)\operatorname{var}(x_1) - \operatorname{cov}(x_2, y)\operatorname{cov}(x_1, x_2)}{\operatorname{var}(x_1)\operatorname{var}(x_2) - [\operatorname{cov}(x_1, x_2)]^2}$$

There is another intuitive way to approach this question. To start, we need to know that the regression line goes through the sample mean $(\overline{x}_1, \overline{x}_2, \overline{y})$. This is easy to verify, by substituting values for $(\overline{x}_1, \overline{x}_2, \overline{y})$ into *any* regression line and show that it satisfies the equality. This is a helpful lemma because we can now run the following regression:

$$\tilde{y}_i \quad = \quad \beta_2 \tilde{x}_{1i} + \beta_3 \tilde{x}_{2i} + u_i$$

where $\tilde{x}_{1i} = x_{1i} - \overline{x}_1$, $\tilde{x}_{2i} = x_{2i} - \overline{x}_2$ and $\tilde{y}_i = y_i - \overline{y}$. In fact, instead of regressing y_i on an intercept, x_1 and x_2, we regress $y_i - \overline{y}$ on $x_{1i} - \overline{x}_1$ and $x_{2i} - \overline{x}_2$ (with no intercept) which is to say, we de-mean the data (subtract average values of each variable) before running the regression. As a result, this regression goes through the origin, and has no intercept which helps us to reduce one parameter from the model and solve the foc system only for two parameters. This transformation leaves the slope parameters intact:

$$\begin{aligned}
\partial RSS/\partial\beta_2\big|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} &= -2\sum_{i\in N} \tilde{x}_{1i}\left(\tilde{y}_i + \widehat{\beta}_2\tilde{x}_{1i} + \widehat{\beta}_3\tilde{x}_{2i}\right) \\
\partial RSS/\partial\beta_3\big|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} &= -2\sum_{i\in N} \tilde{x}_{2i}\left(\tilde{y}_i + \widehat{\beta}_2\tilde{x}_{1i} + \widehat{\beta}_3\tilde{x}_{2i}\right)
\end{aligned}$$

Re-arrange:

$$\begin{aligned}
0 &= \sum_{i\in N}\tilde{x}_{1i}\tilde{y}_i + \widehat{\beta}_2\sum_{i\in N}\tilde{x}_{1i}^2 + \widehat{\beta}_3\sum_{i\in N}\tilde{x}_{1i}\tilde{x}_{2i} \\
0 &= \sum_{i\in N}\tilde{x}_{2i}\tilde{y}_i + \widehat{\beta}_2\sum_{i\in N}\tilde{x}_{2i}\tilde{x}_{1i} + \widehat{\beta}_3\sum_{i\in N}\tilde{x}_{2i}^2
\end{aligned}$$

Re-arrange:

$$\begin{aligned}
0 &= \operatorname{cov}(\tilde{x}_1, \tilde{y}) + \widehat{\beta}_2\operatorname{var}(\tilde{x}_1) + \widehat{\beta}_3\operatorname{cov}(\tilde{x}_1, \tilde{x}_2) \\
0 &= \operatorname{cov}(\tilde{x}_2, \tilde{y}) + \widehat{\beta}_2\operatorname{cov}(\tilde{x}_2, \tilde{x}_1) + \widehat{\beta}_3\operatorname{var}(\tilde{x}_2)
\end{aligned}$$

Re-write in terms of $\widehat{\beta}_2$:

$$\begin{aligned}
\widehat{\beta}_2 &= -(\operatorname{cov}(\tilde{x}_1, \tilde{y}) + \widehat{\beta}_3\operatorname{cov}(\tilde{x}_1, \tilde{x}_{2i}))/\operatorname{var}(\tilde{x}_1) \\
\widehat{\beta}_2 &= -(\operatorname{cov}(\tilde{x}_2, \tilde{y}) + \widehat{\beta}_3\operatorname{var}(\tilde{x}_2))/\operatorname{cov}(\tilde{x}_2, \tilde{x}_1)
\end{aligned}$$

Equate the right hand sides:

$$\frac{\operatorname{cov}(\tilde{x}_1, \tilde{y}) + \widehat{\beta}_3\operatorname{cov}(\tilde{x}_1, \tilde{x}_2)}{\operatorname{var}(\tilde{x}_1)} \quad = \quad \frac{\operatorname{cov}(\tilde{x}_2, \tilde{y}) + \widehat{\beta}_3\operatorname{var}(\tilde{x}_2)}{\operatorname{cov}(\tilde{x}_2, \tilde{x}_1)}$$

Therefore:

$$\operatorname{cov}(\tilde{x}_2, \tilde{x}_1)\operatorname{cov}(\tilde{x}_1, \tilde{y}) + \widehat{\beta}_3[\operatorname{cov}(\tilde{x}_2, \tilde{x}_1)]^2 \quad = \quad \operatorname{var}(\tilde{x}_1)\operatorname{cov}(\tilde{x}_2, \tilde{y}) + \widehat{\beta}_3\operatorname{var}(\tilde{x}_1)\operatorname{var}(\tilde{x}_2)$$

Therefore:

$$\begin{aligned}
\widehat{\beta}_2 &= \frac{\operatorname{var}(\tilde{x}_2)\operatorname{cov}(\tilde{x}_1, \tilde{y}) - \operatorname{cov}(\tilde{x}_1, \tilde{x}_2)\operatorname{cov}(\tilde{x}_2, \tilde{y})}{\operatorname{var}(\tilde{x}_2)\operatorname{var}(\tilde{x}_1) - [\operatorname{cov}(\tilde{x}_2, \tilde{x}_1)]^2} \\
\widehat{\beta}_3 &= \frac{\operatorname{var}(\tilde{x}_1)\operatorname{cov}(\tilde{x}_2, \tilde{y}) - \operatorname{cov}(\tilde{x}_2, \tilde{x}_1)\operatorname{cov}(\tilde{x}_1, \tilde{y})}{\operatorname{var}(\tilde{x}_1)\operatorname{var}(\tilde{x}_2) - [\operatorname{cov}(\tilde{x}_2, \tilde{x}_1)]^2}
\end{aligned}$$

which are the slope estimators. We need to complete one more step, which is to find the original regression intercept. However, once we have $\widehat{\beta}_2$ and $\widehat{\beta}_3$ then we can use them to uncover the intercept: $\widehat{\beta}_1 = \overline{Y} - \widehat{\beta}_2\overline{x}_1 - \widehat{\beta}_3\overline{x}_2$.

4.2 We use the population model $\tilde{y}_i = \beta_2 \tilde{x}_{1i} + \beta_3 \tilde{x}_{2i} + u_i$ together with the following:

$$\begin{aligned}
\mathbb{E}\widehat{\beta}_2 &= \frac{\operatorname{var}(\tilde{x}_2)\operatorname{cov}(\tilde{x}_1,\tilde{y}) - \operatorname{cov}(\tilde{x}_1,\tilde{x}_2)\operatorname{cov}(\tilde{x}_2,\tilde{y})}{\operatorname{var}(\tilde{x}_2)\operatorname{var}(\tilde{x}_1) - [\operatorname{cov}(\tilde{x}_2,\tilde{x}_1)]^2} \\
\mathbb{E}\widehat{\beta}_3 &= \frac{\operatorname{var}(\tilde{x}_1)\operatorname{cov}(\tilde{x}_2,\tilde{y}) - \operatorname{cov}(\tilde{x}_2,\tilde{x}_1)\operatorname{cov}(\tilde{x}_1,\tilde{y})}{\operatorname{var}(\tilde{x}_1)\operatorname{var}(\tilde{x}_2) - [\operatorname{cov}(\tilde{x}_2,\tilde{x}_1)]^2}
\end{aligned}$$

Noting that by (strict) exogeneity assumption $\operatorname{cov}(\tilde{x}_1, u) = 0$ and $\operatorname{cov}(\tilde{x}_2, u) = 0$ then:

$$\begin{aligned}
\mathbb{E}\widehat{\beta}_2 &= \frac{\operatorname{var}(\tilde{x}_2)\operatorname{cov}(\tilde{x}_1,\beta_2\tilde{x}_1+\beta_3\tilde{x}_2+u) - \operatorname{cov}(\tilde{x}_1,\tilde{x}_2)\operatorname{cov}(\tilde{x}_2,\beta_2\tilde{x}_1+\beta_3\tilde{x}_2+u)}{\operatorname{var}(\tilde{x}_2)\operatorname{var}(\tilde{x}_1) - [\operatorname{cov}(\tilde{x}_2,\tilde{x}_1)]^2} \\
&= \frac{\beta_2\operatorname{var}(\tilde{x}_2)\operatorname{var}(\tilde{x}_1) + \beta_3\operatorname{var}(\tilde{x}_2)\operatorname{cov}(\tilde{x}_1,\tilde{x}_2) - \beta_2[\operatorname{cov}(\tilde{x}_1,\tilde{x}_2)]^2 - \beta_3\operatorname{cov}(\tilde{x}_1,\tilde{x}_2)\operatorname{var}(\tilde{x}_2)}{\operatorname{var}(\tilde{x}_2)\operatorname{var}(\tilde{x}_1) - [\operatorname{cov}(\tilde{x}_2,\tilde{x}_1)]^2} \\
&= \frac{\beta_2\{\operatorname{var}(\tilde{x}_2)\operatorname{var}(\tilde{x}_1) - [\operatorname{cov}(\tilde{x}_1,\tilde{x}_2)]^2\} + \beta_3\{\operatorname{var}(\tilde{x}_2)\operatorname{cov}(\tilde{x}_1,\tilde{x}_2) - \operatorname{cov}(\tilde{x}_1,\tilde{x}_2)\operatorname{var}(\tilde{x}_2)\}}{\operatorname{var}(\tilde{x}_2)\operatorname{var}(\tilde{x}_1) - [\operatorname{cov}(\tilde{x}_2,\tilde{x}_1)]^2} \\
&= \beta_2
\end{aligned}$$

The similar derivation holds for $\mathbb{E}\widehat{\beta}_3 = \beta_3$.