

FINANCIAL ECONOMETRICS
SHANGHAI JIAO TONG UNIVERSITY
SUMMER SCHOOL – 2025
EMPIRICAL EXERCISE – 2A –
With Solutions

FATIH KANSOY

Reference: Stock, J.H. and Watson, M.W., Introduction to Econometrics (4th Edition, Global Edition)

Data: Growth.csv from <http://www.pearsonglobaleditions.com>

Empirical Exercise 4.1. On the text website, you will find the data file **Growth**, which contains data on average growth rates from 1960 through 1995 for 65 countries, along with variables that are potentially related to growth. A detailed description is given in **Growth_Description**, also available on the website. In this exercise, you will investigate the relationship between growth and trade.

- (a) Construct a scatterplot of average annual growth rate (*Growth*) on the average trade share (*TradeShare*). Does there appear to be a relationship between the variables?
- (b) One country, Malta, has a trade share much larger than the other countries. Find Malta on the scatterplot. Does Malta look like an outlier?
- (c) Using all observations, run a regression of *Growth* on *TradeShare*. What is the estimated slope? What is the estimated intercept? Use the regression to predict the growth rate for a country with a trade share of 0.5 and for another with a trade share equal to 1.0.
- (d) Estimate the same regression, excluding the data from Malta. Answer the same questions in (c).

- (e) Plot the estimated regression functions from (c) and (d). Using the scatterplot in (a), explain why the regression function that includes Malta is steeper than the regression function that excludes Malta.
- (f) Where is Malta? Why is the Malta trade share so large? Should Malta be included or excluded from the analysis?

Solution:

Data Overview

The dataset contains cross-country data for 65 countries covering the period 1960–1995:

Variable	Description	Mean	Std. Dev.
growth	Average annual GDP growth rate (%)	1.94	1.90
tradeshare	(Exports + Imports) / GDP	0.56	0.29
country_name	Country identifier	—	—

Note: Trade share values greater than 1.0 indicate that the sum of exports and imports exceeds GDP, which can occur in small, open economies that serve as trade hubs.

Key Notation and Formulas

The simple linear regression model is:

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

The OLS estimators are:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

The predicted value for observation i is:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

(a) Scatterplot: Growth vs Trade Share

[See Figure: Scatterplot of Growth vs TradeShare]

Visual Inspection:

The scatterplot reveals a **weak positive relationship** between trade share and economic growth. Countries with higher trade shares tend to have somewhat higher growth rates, though there is considerable scatter around this trend.

Key Observations:

- Most countries cluster with trade shares between 0.2 and 0.8
- Growth rates range from approximately -3% to $+7\%$
- One country (Malta) stands out with an unusually high trade share (≈ 2.0)
- The relationship, if any, appears modest—trade share alone does not strongly predict growth

(b) Identifying Malta as an Outlier

Countries with Trade Share > 1.0 :

Country	Trade Share	Growth (%)
Malaysia	1.11	4.11
Belgium	1.12	2.65
Mauritius	1.13	3.02
Malta	1.99	6.65

Trade Share Distribution:

	Mean	Std. Dev.	Median	95th Pctl	Max
Trade Share	0.56	0.29	0.54	1.11	1.99

Analysis:

Malta is clearly an **outlier in the X variable** (trade share). Its trade share of 1.99 is:

- Nearly **twice the 95th percentile** (1.11)
- More than **3.5 standard deviations** above the mean: $(1.99 - 0.56)/0.29 = 4.9$
- **Far removed** from all other observations in the horizontal direction

In regression terminology, Malta is a **high-leverage point**—an observation with an extreme X value that can have disproportionate influence on the estimated slope.

(c) Regression: All Observations (Including Malta)

Model: $\text{Growth}_i = \beta_0 + \beta_1 \times \text{TradeShare}_i + u_i$

Regression Output

	Coefficient	Std. Error	t-statistic	p-value
TradeShare	2.306	0.773	2.98	0.004
Constant	0.640	0.490	1.31	0.196
$n = 65, \quad R^2 = 0.124, \quad \text{Root MSE} = 1.79$				

Estimated Regression Equation

$$\widehat{\text{Growth}} = 0.64 + 2.31 \times \text{TradeShare}$$

Interpretation:

- **Slope** ($\hat{\beta}_1 = 2.31$): A one-unit increase in trade share is associated with a 2.31 percentage point increase in the average annual growth rate.
- **Intercept** ($\hat{\beta}_0 = 0.64$): A country with zero trade share would have a predicted growth rate of 0.64%. (This is an extrapolation—no country has zero trade.)
- $R^2 = 0.124$: Trade share explains 12.4% of the variation in growth rates.

Predictions

Using the formula $\widehat{\text{Growth}} = 0.64 + 2.31 \times \text{TradeShare}$:

Trade Share	Predicted Growth Rate
0.5	$0.64 + 2.31 \times 0.5 = \mathbf{1.79\%}$
1.0	$0.64 + 2.31 \times 1.0 = \mathbf{2.95\%}$

(d) Regression: Excluding Malta

Regression Output

	Coefficient	Std. Error	t-statistic	p-value
TradeShare	1.681	0.987	1.70	0.094
Constant	0.957	0.580	1.65	0.104
$n = 64,$ $R^2 = 0.045,$ Root MSE = 1.79				

Estimated Regression Equation

$$\widehat{\text{Growth}} = 0.96 + 1.68 \times \text{TradeShare}$$

Predictions

Trade Share	Predicted Growth Rate
0.5	$0.96 + 1.68 \times 0.5 = \mathbf{1.80\%}$
1.0	$0.96 + 1.68 \times 1.0 = \mathbf{2.64\%}$

Comparison: With vs Without Malta

	With Malta ($n = 65$)	Without Malta ($n = 64$)
Intercept ($\hat{\beta}_0$)	0.64	0.96
Slope ($\hat{\beta}_1$)	2.31	1.68
R^2	0.124	0.045
Predicted Growth (TS = 0.5)	1.79%	1.80%
Predicted Growth (TS = 1.0)	2.95%	2.64%

Key Finding: Removing Malta reduces the slope by 27% (from 2.31 to 1.68) and dramatically lowers R^2 from 12.4% to just 4.5%.

(e) Comparing the Regression Functions

[See Figure: Both Regression Lines Plotted]

Why is the regression line steeper when Malta is included?

1. **Leverage:** Malta has an extreme X value (trade share = 1.99), placing it far from the mean of X ($\bar{X} = 0.56$). In OLS regression, observations far from $\bar{X}$ have high *leverage*—they exert disproportionate “pull” on the fitted line.
2. **Position:** Malta lies *above* the regression line that would be fitted without it. The combination of high leverage and being above the line means Malta pulls the regression line upward at high trade share values.
3. **Mechanical effect:** Including Malta increases $\hat{\beta}_1$ from 1.68 to 2.31 because the OLS estimator is:

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

Malta contributes a large positive term to the numerator (since both $(X_{\text{Malta}} - \bar{X})$ and $(Y_{\text{Malta}} - \bar{Y})$ are positive and large).

Summary: Malta is a **high-leverage, influential observation**. Its extreme trade share and above-average growth rate combine to steepen the estimated relationship between trade and growth.

(f) Should Malta Be Included or Excluded?

Where is Malta?

Malta is a small island nation in the Mediterranean Sea, located south of Sicily (Italy). With a population of approximately 500,000, it is one of the world's smallest and most densely populated countries.

Why is Malta's Trade Share So Large?

Malta's trade share of approximately 2.0 (meaning imports plus exports are twice its GDP) reflects its unique economic role:

- **Transshipment hub:** Malta serves as a major freight transport and logistics center. Goods arrive (counted as imports) and are immediately re-exported to other countries (counted as exports).
- **Re-export economy:** Unlike most countries where imports are consumed domestically and exports are domestically produced, Malta's trade figures include substantial "pass-through" trade.
- **Small domestic economy:** With a tiny domestic market, even modest absolute trade volumes translate to large trade-to-GDP ratios.

Should Malta Be Included?

Arguments for exclusion:

- Malta's "trade" is fundamentally different from that of typical countries—it measures transshipment activity, not economic openness in the usual sense.
- Including Malta may distort our understanding of the trade-growth relationship for typical economies.
- Malta is a statistical outlier that exerts undue influence on the regression results.

Arguments for inclusion:

- There is no data error—Malta's trade share is correctly measured.
- Excluding observations simply because they are influential can bias results.

Recommendation: Malta should likely be **excluded** from this analysis. Its trade share reflects transshipment activity rather than the economic openness that the trade-growth hypothesis concerns. For understanding the relationship between trade openness and growth in typical economies, the regression excluding Malta ($\hat{\beta}_1 = 1.68$) is more appropriate.

Summary: Key Takeaways

Question	Finding	Interpretation
Is there a relationship?	Yes, weak positive	Higher trade $\rightarrow$ higher growth
Is Malta an outlier?	Yes (trade share = 1.99)	High-leverage observation
Slope with Malta	2.31	Overstates relationship
Slope without Malta	1.68	More representative
Include Malta?	No	Transshipment $\neq$ openness

The Big Picture

This exercise illustrates several important econometric concepts:

1. **Outliers and Influence:** A single observation can dramatically affect regression results. Malta changes the slope by 27% and the R^2 by nearly three-fold.
2. **Leverage:** Observations with extreme X values have high leverage and can “pull” the regression line toward themselves.
3. **Economic vs. Statistical Outliers:** Malta is not a data error—it is a country with genuinely unusual characteristics. The decision to include or exclude it should be based on economic reasoning (is Malta’s trade comparable to other countries?), not just statistical criteria.
4. **Robustness:** When a single observation substantially changes results, conclusions should be reported both with and without that observation, along with economic justification for the preferred specification.

Calculations performed using Stata. See accompanying `EE_2A_solution.do` file.