

# Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

<b>Student:</b>	<b>Bc. Rufat Asadli</b>
<b>Advisor:</b>	<b>RNDr. Michal Červinka, Ph.D.</b>
<b>Title of the thesis:</b>	<b>Analysis of Czech Trade Structure Using the Zipf's Law</b>

## **OVERALL ASSESSMENT** (provided in English, Czech, or Slovak):

### **Short summary**

Revealed comparative advantage (RCA) is examined for the Czech Republic. The adherence of this data to Zipf's law is tested. Zipf's law states that there is a negative linear relationship between the data and the rank of the data.

The analysis is split into two main parts:

1/ Pooled OLS regression to verify that Zipf's law holds for the RCA panel data.  
 $\log(\text{RCA}-0.5) \sim \log(\text{rank}(\text{RCA}))$  regression - the slope is hypothesised to be -1.

This slope, 'exponent', is also estimated using OLS for each year separately. This time series is then temporally disaggregated to obtain quarterly data that is to be used in the second part.

2/ The second part of the empirical section uses this estimated 'exponent' time series from the first part, and it explores whether there is some relationship with macroeconomic variables (GDP, FDI, inflation and Real effective exchange rate). Three different specifications of the VAR model are used, one of them being Bayesian VAR. Here it is concluded that an economic shock might cause a reduction in comparative advantage.

### **Contribution**

The author explores a topic which has been, to some extent, covered in the previous literature. The paper with a similar research question by Hinloopen and van Marrewijk (2009) explores an older dataset from 1970 to 1997. The data used in this thesis covers trade development from 1998 to 2020 in the Czech Republic. The use of an up-to-date dataset with a focus on one country is novel.

### **Methods**

OLS and pooled OLS regression are used in the first part. Vector autoregressive model (VAR) and Bayesian VAR (BVAR) are used in the second part of the estimation.

The models and methods used are advanced for a bachelor's degree student. Unfortunately, from my point of view, the thesis is missing a coherent and organised methodology section that would (together with interpretation) show the understanding of the methods being applied. The methods used are discussed in multiple sections (i.e. VAR in 5.2, 5.3 and 6.2) with varying levels of detail. I would also appreciate more details about BVAR.

It is nice that the R code with the data is provided. It is pretty organised and corresponds to the reported results.

### **Literature**

The literature review section overviews relevant literature about trade in the Czech Republic, revealed comparative advantage and Zipf's law.

### **Manuscript form**

The organisation and structure of the thesis, its sections, and paragraphs could be greatly improved. The data, methodology and empirical results sections are the most disorganised and confusing parts. Sentences are often hard to follow and understand, and in need of proofreading. Sentence clarity could be improved, i.e. by using simpler or shorter sentences. The thesis is typeset in LaTeX, and it is referenced properly.

# Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

<b>Student:</b>	<b>Bc. Rufat Asadli</b>
<b>Advisor:</b>	<b>RNDr. Michal Červinka, Ph.D.</b>
<b>Title of the thesis:</b>	<b>Analysis of Czech Trade Structure Using the Zipf's Law</b>

More specific comments:

- From the data section (5.1), it is not clear what variables you are working with, at which frequency, for what period and how (or if) they are preprocessed (logs or other transformations, dropping observations based on some rule). The use of summarising tables might be useful to the reader because as is, this information is scattered around in the thesis.
- Interpreting the results (Table 3) before we even know what data went into the model. Log returns, logs or double differenced data are used, but this is only mentioned at the end of this section. ADF test results for these transformed variables are not mentioned.
- Contradicting information about the models and data used. For example:
  - o 'three different models with three consecutive lags.' when we have models with 3,4 and 5 lags
  - o or 'The former VAR(3) takes log-transformed data' while this model used log returns.

More minor problems:

- Thesis is not very well interconnected. References to different sections of the thesis were not carefully checked. i.e. 'This model is previously defined Vector Autoregression (VAR)' – the model was not previously defined.
- 'notable' or interesting things are mentioned, but the reason why are they notable is not mentioned.
- Sometimes, words with a different meaning than intended were used
- Tables should be self-explanatory (what numbers are shown in parenthesis, confidence bands not labelled, etc.).
- In some cases, the abbreviation was not appropriately introduced (i.e. H-O model)

## Suggested questions for the discussion during the defense:

1/ Data preprocessing:

- We are dropping RCA observations if they are lower than 1. Shouldn't dropped observations also follow Zipf's law? Why are we only interested in exports where we have a comparative advantage?
- Discuss different ways how to scale and transform the FDI variable (which is nonstationary, and contains negative values) to achieve stationarity. For example, scaling the data to be positive and taking logs or log returns or shifting the data by constant to get positive values and then taking log returns. Compare with the approaches used (dropping negative values and taking log returns, double differencing).

2/ Methodology:

- Discuss the motivation behind using a VAR model instead of a univariate time series model.
- Discuss more advantages and disadvantages of BVAR. Can it be used in the same way as VAR? What about standard errors of BVAR?
- How, why and when should we test for cointegration of time series?
- The autocorrelation of variables (Figure 8, diagonal plots) does not seem small – what do the numbers mean, and how can it be interpreted?

3/ Author's main results and conclusions from them:

- 'the Zipf's Law hypothesizes that power law exponent  $\alpha$  is equal to 1.' - In the first model (OLS and pooled OLS), how would you formally test this?
- The second model: Interpret results for each model in Table 3 and note differences between different VAR specifications (data, units, preprocessing,...).
- Interpret impulse responses from the BVAR model (Figure 7), from which you conclude that an economic shock may cause a reduction of comparative advantage. What are the units, and are the results statistically significant? How are confidence bands for BVAR calculated?

# Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	Bc. Rufat Asadli
Advisor:	RNDr. Michal Červinka, Ph.D.
Title of the thesis:	Analysis of Czech Trade Structure Using the Zipf's Law

In my view, the thesis fulfils the requirements for a bachelor thesis at IES, Faculty of Social Sciences, Charles University; That is, I suggest grade C, assuming the student can adequately answer the above questions during the defense.

The results of the Urkund analysis do not indicate significant text similarity with other available sources.

**SUMMARY OF POINTS AWARDED** (for details, see below):

CATEGORY	POINTS
Contribution (max. 30 points)	25
Methods (max. 30 points)	16
Literature (max. 20 points)	20
Manuscript Form (max. 20 points)	10
<b>TOTAL POINTS</b> (max. 100 points)	<b>71</b>
<b>GRADE</b> (A – B – C – D – E – F)	<b>C</b>

**NAME OF THE REFEREE:** Lenka Nechvátalová

**DATE OF EVALUATION:** 27.5.2022

**Digitally signed (27.5.2022):**  
Lenka Nechvátalová

---

**Referee Signature**

**EXPLANATION OF CATEGORIES AND SCALE:**

**CONTRIBUTION:** *The author presents original ideas on the topic demonstrating critical thinking and ability to draw conclusions based on the knowledge of relevant theory and empirics. There is a distinct value added of the thesis.*

**METHODS:** *The tools used are relevant to the research question being investigated, and adequate to the author's level of studies. The thesis topic is comprehensively analyzed.*

**LITERATURE REVIEW:** *The thesis demonstrates author's full understanding and command of recent literature. The author quotes relevant literature in a proper way.*

**MANUSCRIPT FORM:** *The thesis is well structured. The student uses appropriate language and style, including academic format for graphs and tables. The text effectively refers to graphs and tables and disposes with a complete bibliography.*

**Overall grading:**

TOTAL	GRADE
91 – 100	A
81 - 90	B
71 - 80	C
61 – 70	D
51 – 60	E
0 – 50	F