CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Impact of Digital Service Tax on tax revenues of EU members

Bachelor's thesis

Author: František Kosař
Study program: Economics and Finance
Supervisor: doc. Petr Janský, Ph.D.
Year of defense: 2022

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, August 2, 2022

František Kosař

Abstract

This thesis is to dive into a implementation of Digital Service Tax in the European Union and to approximate its influence on tax revenues of respective EU members. The purpose of this work is to help understand corporation behaviour with respect to taxation in the EU region. This approximation was carried out using data from institutions: Eurostat, OECD and TaxFoundation.org. The analysis consist of descriptive statistics of respective variables and utilisation of panel data regression methods (fixed-effect and first-difference). The analysis arrives at discovery, that at statistical significance level of 15%, there is a positive relationship between presence of digital service tax tool and the corporate tax revenue collected.

JEL Classification	F12, F21, F23, H25, H71, H87
Keywords	economics, tax, multinational enterprises, European Union

Title	Impact of Digital Service Tax on tax revenues of EU members

Author's e-mail	79260198@fsv.cuni.cz
Supervisor's e-mail	petr.jansky@fsv.cuni.cz

Abstrakt

Tato práce se zabývá implementací Digitální daně v Evropské Unii a snaží se odhadnout její vliv na rozpočty jendotlivý členských státu. Účel této práce je snaha pochopit chování korporátní chování vzhledem k daňový povinnostem v evropském regionu. Tato aproximace byla provedena za využití dat z institucí: Eurostat, OECD, TaxFoundation.org. Analýza se skládá z dekriptivní statistiky pro jednotlivé proměnné a z využití regresních metod pro panelová data (fixed-effect a first-difference). Analýza dochází ke zjištění, že na úrovni signifikance 15% existuje pozitivní vztah mezi prezencí digitální daně a vybranou korporátní daní.

Klasifikace JEL F12, F21, F23, H25, H71, H87
Klíčová slova ekonomie, daň, mezinárodní společnosti, Evropská Unie

Název práce Vliv digitální daně v EU na rozpocty clenskych statu

E-mail autora 79260198@fsv.cuni.cz
E-mail vedoucího práce petr.jansky@fsv.cuni.cz

Acknowledgments

The author is grateful especially to doc. Petr Janský, Ph.D., for his contribution and insightful comments on my thesis.

Typeset in FSV LaTeX template with great thanks to prof. Zuzana Havrankova and prof. Tomas Havranek of Institute of Economic Studies, Faculty of Social Sciences, Charles University.

Bibliographic Record

Kosař, František: *Impact of Digital Service Tax on tax revenues of EU members.* Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2022, pages 43. Advisor: doc. Petr Janský, Ph.D.

Contents

List of Tables

List of Figures

Acronyms

DTS Digital Service Tax

MNE Multinational Enterprises

MNC Multinational Corporation

BEPS Base Erosion and Profit Shifting

BEAT Base Erosion Anti-abuse

METR Minimal Effective Tax Rate

DESI Digital Economy and Society Index

EU European Union

ITCI International Tax Competitiveness Index

OLS Ordinary Least Squares

OECD Organisation for Economic Co-operation and Development

SD Standard Deviation

IIR Income Inclusion Rule

UTPR Undertaxed Payment Rule

STTR Subject to Tax Rule

Bachelor's Thesis Proposal

Author	František Kosař
Supervisor	doc. Petr Janský, Ph.D.
Proposed topic	Impact of Digital Service Tax on tax revenues of EU members

Motivation My research should focus on approximating the effect of Digital Service Tax on revenues of respective governments in the EU and also discuss how the BigTech corporations would behave under such tax laws in the European market.

The progressive nature of multinational corporations in digital industry and its outstanding ability to use BEPS (Base Erosion and Profit Shifting as defined by OECD) tools enables Bigtech corporations to engage in illicit financial operations and as a result total revenue loss due to BEPS is approximated to be close to US$ 1 trillion per year (Garcia-Bernardo and JanskĂ˝, 2021). Hence this topic is very important to study in order to unveil the effect and create arguments for or against such tax legislation.

Methodology Assuming the digital service tax as proposed by EU in 2018 (Proposal for a Council Directive, European Commission, 2018), I plan on using the data provided by the European Union containing information about corporate taxes in past years to analyze the shift of revenue from big digital companies in countries where some form of digital service tax is already in place (Tax Foundation, Elke Asen, 2021) and build a fitting model around it. Model should simulate the relationship between total tax revenue gain and variables such as population, GDP, digitalisation of the market and other characteristics of respective jurisdictions. Then with some alternations I would try to expand this model also to other countries in EU and hence approximate the total revenue, that could be raised by such a tool as digital service tax. Furthermore, I will use data from big corporations participating in country-by-country reporting initiative (OECD.Stat database) and accounting details from EU company registers in order to train my model and predict the behaviour of other companies in the same industry, which would help me to design a fitting predictive model.

Core bibliography

Garcia-Bernando J. JanskĂ˝ P. (2021), Profit shifting of multinational corporations worldwide

Bauer M (2018), Five questions about the digital services tax to Pierre Moscovici

Cobham A. JanskĂ˝ P. (2020), Estimating Illicit Financial Flows

Garcia-Bernardo, JanskĂ˝ P. Torslov T. (2020), Multinational corporations and tax havens: evidence from country-by-country reporting

European Commission (2018), Proposal for a council directive laying down rules relating to the corporate taxation of a significant digital presence

Author

Supervisor

Chapter 1

Introduction

All throughout history, authorities in form of chieftains, kings or governments needed to raise funds in order to finance their expenditures. The oldest known form of what could be called a tax, comes from ancient Egypt, where peasants would give a fifth of their harvest to the Pharaon (Burg (2004)). Early taxes were used to build armies, support the ruling class and build defences. Today taxes are used in a more diverse manner and governments can provide a wide range of services to the public. However, with the rise of humanity and economic systems also legislation and tax systems, in particular, had to reflect this change. Hence today economic agents all over the world can be subjected to more specific and efficient taxing tools such as income tax, corporate tax, capital gain tax, inheritance tax, property tax or sales tax.

Tax revenue is an essential part of the governmental budget in almost every country in the world and enables governments to provide for their subjects. According to the World Bank, tax revenues above 15% of a countryâ€™s gross domestic product (GDP) are a key ingredient for economic growth and, ultimately, poverty reduction (Corbo *et al.* (1991)). For instance, from OECD data about national taxes average rate of tax revenue to GDP is 33,8% (Data (2018))

In recent years countries all over the world have been debating significant changes to tax systems and rules that apply to multinational companies. This initiative responds to concerns raised by politicians and economists alike, that the tax systems tend not to reflect the digitalisation of the economy. A special part of these companies is big digital technological companies (â€žBigTechâ€ś). Traditionally, firms pay the majority of their taxes in places of production, but if we are talking about digital firms, that sell a digital product it is generally rather complicated to design a legislative framework that connects a product to a consumer and hence it is easy for digital firms to shift their profits to regions imposing lower tax rates on corporates residing within. These practices are not always illegal. However, apart from lost tax revenue, the profit shifting also gives advantage to the big digital companies. Through such measures, the big digital gain competitive advantages that further endanger domestic tax base. These factors contributed to the proposition of the European Commission to introduce a new tax rule in 2018, which was aimed to stop multinational enterprises (MNE) in their efforts to shift profits to jurisdictions with lower effective tax rates through digital channels. These new rules are part of a larger global initiative of OECD/G20 to address Base Erosion and Profit Shifting of multinational enterprises. In theory, this tool is supposed to lower the amount of total tax revenue loss, which is estimated to reach over US$ 1 trillion dollars per year (Jansky *et al.* (2020)).

In this paper, we would like to dive deep into the analyses of the Corporate Tax Revenues in Europe and to approximate an influence of Digital Service Tax on collected tax revenue. Furthermore we would like to inspect additional factor, that might influence the Corporate Tax Revenue and incorporate them into the model. The model should uncover whether a digital service tax would improve tax codes of selected countries or whether the digital companies would be incentivised to shift their profits to lower-taxed jurisdictions.

Chapter 2

Literature Review

In this chapter we would like to go through know research and papers that are directed at the digital service tax and at international taxation initiatives. Furthermore we would like to go trough most recent and essential legislation proposals, that proposes either an implementation of a digital service tax tool or a substitute to such a tool.

2.1 Digital Service Tax Legislation

Since 2019 countries all over the Europe started to implement or propose their own version of a taxation of digital economy. Till this day, 15 European countries have implemented or proposed their version of a digital service tax. The key initiative came in 2013 from Organization for Economic Co-operation and Development (OECD). This initiative can be divided into two so-called Pillars, that are designed to address challenges arising from the digitalisation of the economy.(Popescu (2020)).

2.1.1 BEPS - Pillar One

Pillar One is a piece of multinational legislation that is agreed upon by more than 130 countries of the world and it focuses on a question "Where should the taxes be paid". The Pillar One is to be applied to large multinational corporations (MNC) and is to reallocate some taxable income of MNCs to market jurisdictions. Originally the proposal was aimed to tackle profit shifting and base erosion connected to highly digitalised companies such as Google or Apple, however since then the scope has moved far from the initial intentions. The expectation is that the rules will be finalized by 2022 with an effective date in 2023. (Popescu (2020)).The key point to take away from this proposed piece of legislation are that profits should be taxed in the end-market jurisdiction where goods and services are used or consumed, it will apply to groups with 20 billion euro in worldwide revenues and that Digital services taxes and other similar measures will be repealed once the Pillar One legislation is effective. Some aspects of the proposal are still being discussed (Popescu (2020))

2.1.2 BEPS - Pillar Two

Pillar Two is also a piece of multinational legislation that is agreed upon by more than 130 countries of the world and it focuses on a question "How high tax rate should be imposed on MNE". The Pillar Two aims to set an appropriate tax rate and introduces a set of rules, that are meant to ensure, that the taxes will be paid. In 2021 OECD/G20 nclusive Framework (IF) on Base Erosion and Profit Shifting (BEPS) released above mentioned set of Model Rules. These rules represent "common global approach" and the European Union proposed a directive to incorporate these rules into the EU law. These rules set a Global Minimum Tax at 15% for multinational Enterprises with turnover of more than 750 million EUR. Pillar Two has three new rules granting jurisdictions additional taxing rights, including: â€” Two interlocking domestic rules (the GloBE rules) that are the subject of the Model Rules:

1. An Income Inclusion Rule (IIR), which imposes top-up tax on a parent entity in respect of the income of subsidiaries and permanent establishments that is taxed at less than a 15% minimum effective tax rate

2. A supporting Under-taxed Payment Rule (UTPR), which denies deductions or requires an equivalent adjustment in the event a parent entityâ€™s allocate able share of the top-up tax regarding a low taxed constituent entity is not subject to tax under an IIR

3. A Subject to Tax Rule (STTR), which overrides treaty benefits for certain related-party payments (including interest and royalties) that are not subject to a 9% minimum rate of tax in the recipient jurisdiction. The STTR will be creditable as a covered tax under the IIR and UTPR, i.e., the STTR applies first.

Chapter 3

Data

3.1 Data Introduction

The data for this thesis come from several different sources and were compiled together using functions in R software. Among main sources are OECD open database (Data (2018)), Tax Foundation (Bunn & Asen (2020)) or Eurostat(Russo (2020)) Altogether we were able to comprise 90 observations from 18 countries. The observations were collected in a time period ranging from 2016 to 2020. We initially acquired data sets that were collected far beyond mentioned time period (2016-2020), however the availability of essential variables caused the data set to shrink. In the following text we would like to further describe respective variables that are included in the data frame. It is worth mentioning that the data set fulfills all important characteristics of panel data (Wooldridge (2015)) and therefore relevant therefore relevant econometric tools should be utilized during the analysis (see *Chapter 4 - Methodology*).

3.2 Variables Description

In this section we would like discuss all data-set-confounding variables with a strong emphasis on the ones that were later used during building the estimator model.

- **Year**: Time variable measured in single years and ranging from 2016 to 2020. This variable is essential for the estimation, since we observe desired phenomenons with distinctive characteristics that are visible in time such as trends or relationships to previous time periods (auto-correlations).

- **Country**: Country of the origin of the observed data. The observations come from total of 18 EU members states. Those are namely: Austria, Belgium, Germany, Denmark, Estonia, Spain, Finland, France, Hungary, Ireland, Italy, Luxembourg, Latvia, Netherlands, Poland, Portugal, Sweden and Slovenia. There are also two additional variables carrying the same information denoted by international country abbreviations, that were used to merge different data frames together. These extra variables will be left in the data set for future work.

- **Corporate tax revenue**: Variable portraying total amount of tax revenue collected from corporate sources. This revenue includes income, profit and capital gain tax of corporate enterprises. The data come from OECD database. The variable is measured in millions of USD.

- **GDP per capita**: Variable chosen to represent economic performance in a region. GDP per capita is computed as total GDP of a country divided by its population. The source of this variable is OECD database. Variable is measured in USD. Among highest values are observations from Luxembourg or Ireland , on the lower end we can see countries such as Estonia, Hungary or Slovenia.

- **The Digital Economy and Society Index** The Digital Economy and Society Index (DESI) is a index published by the European Commission. "DESI summarises indicators on Europe's digital performance and tracks the progress of EU countries." (Bánhidi *et al.* (2020)). The first year that this index research was conducted is 2016.
  The index itself is calculation from 5 respective fields of interest: Connectivity, Digital Skills (also sometimes referred to as Human Capital), Use of Internet, Integration of Digital Technology and Digital Public Services. Each of the mentioned fields is further defined into smaller interest fields and scored according to predefined rules (Bánhidi *et al.* (2020)). Finally each field gets a score ranging from 0 to 1. In can be interpreted such as "the higher the score of a country the more digitalised the country is". The variable we are using is *Total* DESI, that is computed by weighting importance of the results from the sub-fields. The predefined formula to calculate the total index is:

$$DESI_i = Connectivity_i * 0.25 + Human\ Capital_i * 0.25 + Use\ of\ the\ Internet_i * 0.15 + Integration\ of\ Digital\ Technology_i * 0.2 + Digital\ Public\ Services_i * 0.15$$

- **International Tax Competitiveness Index**: International Tax Competitiveness Index (ITCI) is a metric published by an international organisation called Tax Foundation.org. The index itself can be interpreted as an amount of "Tax burden" imposed upon economic agents operating on respective country markets. The index depends on tax rates imposed on different parts of economics but also on difficulty for the agents to follow the tax code. The process of deriving the final index is beyond the reach of this paper, however feel free to look deeper into the methodology of the Tax Foundation research team (Bunn & Asen (2020)).
  Theoretically the total ITCI ranges from 0 and 10, where countries with 0 are the least tax competitive and countries with 10 have the most competitive tax code. Among most tax competitive countries in the EU according to the index are Estonia, Latvia and Sweden. The least tax competitive countries are on the other hand France, Portugal or Poland. The variable has no obvious trend and we use it to estimate incentive of digital industry to either stay on the market and report their profits under a tax jurisdiction or to shift their profit to more tax competitive country in order to lower their costs.

- **is.DST**: "is.DST" is a dummy variable signalizing whether there was implemented a Digital Service Tax in particular regions and time periods. The information comes from each respective tax jurisdiction, that implemented a tax that would qualify or is similar enough to DST according to proposed legislative piece by the European Commission (COM (2018)). This paper considers generalised version of DST and even thought the Digital Service Tax tools differ across the EU member states. The whole concept of DST is fairly new and so the the number of observations with present DST is rather low.

Chapter 4

Methodology

4.1 Methodology Introduction

In this section we would like to focus on specific ways on how to analyse the data presented in Chapter 3 and our reasoning, why we chose these methods. Firstly we cleaned and analyzed the data in order to identify distinctive characteristics, that should be accounted for in a model such as trends or mutual relationships between explanatory variables. Secondly we chose and built a relevant linear model and compared their performance against each other. Thirdly we tuned the best performing model in order to reach optimized results. We conducted the whole analysis using R Studio, version 3.6.3.

4.2 Analysing the Data

This section will be similar to Chapter 3, however we would like to dive deeper into our motivation, why we chose previously mentioned variables and what interpretation power could they hold for the thesis.

4.2.1 Merging the data

First of all we had to obtain all the data needed for our research a put them in one unified data set, that would be easy to work with. These separated data sets came from sources mentioned in section 3.2. - Variables Description. We paired the data frames together by matching the year and region of observations. Not all data frames obtained form different sources had a variable documenting the country in the same format. Some saved the country information via 2 or 3 letter (ISO-3166-1 ALPHA-2/3) abbreviation and some had full country name, so we had to use additional help lists that mapped different formatting of country names together.

4.2.2 Clearing the data

Upon merging the data into one compact data set, we selected out data that do not have any value missing, because such observations would be unusable in our linear model.
Then we started to look into the distribution of respective variables and decided whether the variables contain any outliers, that could interfere with our computations. Generally outliers could be defined as "data point that differs significantly from other observations" (Dougherty (2011)). We identified outliers by defining an interval:

$$(Q_1 - 1.5 * IQR; Q_3 + 1.5 * IQR)$$

Where Q_1 and Q_3 stand for values of the lower and upper quartiles respectively and IQR stands for *interquartile range*, hence their difference. When identified, we had to decide, whether the untypical data point was a product of measurement error, human mistake or whether there are significant natural differences in the population. We decided that all the outliers found in the data set are not product of an error and hence we did not exclude them from the final data frame.

4.3 Building a Fixed-Effect Model

As mentioned in Chapter 3, we worked with panel data. Hence we transformed the data by creating unique indexes signalizing the panel structure. The indexes are comprised of year and region of the observation. While building a model we had to establish two main building stones - defining the formula and selecting a correct estimation method.

4.3.1 Deriving the Formula

The formula is designed to reflect relevant factors influencing amount of corporate tax revenue. The initial theoretical formula is:
$Corporate\ Tax\ Revenue_{it} = \beta_0 + \beta_1 * GDP\ per\ * capita_{it} + \beta_2 * DESI_{it} + \beta_3 *$

$ITCI_{it} + \beta_4 * is.DTS_{it} + a_i + u_{it}$

The detailed description of all variables used can be seen in Chapter 3. Briefly: $Corporate\ Tax\ Revenue_{it}$ represents total amount of tax collected by respective states, $GDP\ per\ capita_{it}$ stands for total GDP of a region divided by the population, $DESI_{it}$ signals digitalisation of a region, $ITCI_{it}$ is an International Tax Competitiveness Index and $is.DTS_{it}$ represents a dummy variable corresponding to presence or absence of a Digital Service Tax legislation in a region. The more detailed description is presented in Chapter 3.
Parameter a_i represents a so-called "fixed effect", which can be described as an unobserved time-invariant individual effect, that is distinctive for every region represented in the data set. In our case this unobserved effect can represent cultural differences or specific customer behaviour of the population (Wooldridge (2015)).The final variable is u_{iT} that represent unobservable disturbances that have effect on the estimation (Wooldridge (2015))
Lower indexes i and T represent unique observation and time period respectively.

Upon analysing each variable we have to take into account some specific characteristics about our data. For example variable $DESI_{it}$ is following a trend and hence a new variable should be added to the formula in order to account for such a trend. Otherwise the model could suffer from spurious correlation, that could occur due to more than one variable simply increasing or decreasing in time. Hence this effect could hypothetically violate explanatory variables independence assumption(Wooldridge (2015)). So variable *Year* was added to the formula to account for this characteristic.

Another important assumption is connected to $ITCI_{it}$. This variable is specific in a sense that it takes some time for the market to reflect it. If we put it plainly - when the legislation regarding a tax code changes it is safe to assume that it would change the behaviour of a corporate market agent. Hence we decided to add another variable to the formula representing "lagged" ITCI ($ITCI_{it-1}$).

Due to the fact that variables $Corporate\ Tax\ Revenue_{it}$ and $GDP\ per\ capita_{it}$ are initially given in total amounts of USD, it would be better for us to transform these variables using logarithmic transformation. Firstly we think, that the transformed variable $GDP\ per\ capita_{it}$ better represents the influence on the dependent tax revenue variable but secondly and more importantly, it is easier for us to interpret results emerging out of the model. Using logarithm enables us to view the β_i estimators trough a lens of percentage growth. In other words we are estimating elasticity of the dependent variable.

When adding all relevant variables into the formula, we got a final equation:

$log(Corporate\ Tax\ Revenue_{it}) = \beta_0 + \beta_1 log(GDP\ per\ capita_{it}) + \beta_2 DESI_{it} + \beta_3 ITCI_{it} + \beta_4 ITCI_{it-1} + \beta_5 is.DTS_{it} + \beta_6 Year_i + a_i + u_{it}$

4.3.2 Choosing the Estimation Method

Due to the fact, that we worked with panel data (Chapter 3). We have to account for parameter a_i. There are generally more ways to model panel data, but among the most popular are: first-difference, fixed-effect, random-effect and dummy variable model. We cannot use plain OLS estimators, because of the parameter a_i. With respect to a rather small number of observations N, we decided that fixed-effect model ("within" model) was the most efficient one, provided that all assumption are met. The final equation has to be transformed in order to eliminate influence of the fixed effect a_i (Wooldridge (2015)) The data go though following transformation:

- Lets consider an unobserved effect model
  $y_{it} = \beta_0 + \beta_1 * x_{it} + a_i + u_{it}$ where t = 1 . . . T and n= 1 . . . N
- When unobserved heterogeneity, represented by a_i is correlated with explanatory variable x_{it}, OLS estimates of this model parameters are biased and inconsistent.
- Now we take the unobserved effect model for each i, we average the equation over time:

$$\bar{y_i} = \beta_0 + \beta_1 * \bar{x_i} + a_i + \bar{u_i}$$
$$\text{where } \bar{y_i} = T^{-1} \sum_{t=1}^{T} y_{it}$$

- Subtracting the averages from the original equation, we get time-demeaned model. In this model unobserved effect a_i is eliminated and we are bale to use pooled OLS estimators(Wooldridge (2015)). The time-demeaned model:

$$y_{it} - \bar{y_i} = \beta_0 - \beta_0 + \beta_1 * (x_{it} - \bar{x_i}) + a_i - a_i + u_{it} - \bar{u_i} =$$
$$\ddot{y_{it}} = \beta_1 * x_{it} + \ddot{u_{it}}$$

4.4 Confirming Assumptions

In order to build a well-performing model, relevant assumption about the entry data have to be met. These assumptions are:

1. **Assumption FE1**:
   For each i, the model is

   $y_{it} = \beta_1 x_{it1} + ... + \beta_k x_{itk} + a_i + uit$

   where t=1...T, i=1...n and parameter β_j are to be estimated and a_i is the unobserved fixed effect (unobserved heterogeneity).

2. **Assumption FE2:**
   We have a random sample in the cross-sectional dimension.

3. **Assumption FE3**:
   Each explanatory variable changes over time (for at least some i), and there are no perfect linear relationships among the explanatory variables.

4. **Assumption FE4**:
   For each i, the expected value of the idiosyncratic error given the explanatory variables in all time periods and the unobserved effect is zero: $E(u_{it}|X_i, a_i) = 0$. Under assumptions FE1-FE4, the estimator $\tilde{\beta_{FE}}$ is unbiased. The key assumption is strict exogeneity (FE4). Furthermore the estimators of a fixed-effect model $plm(\tilde{\beta_{FE}}) = \beta$ as $N \to \infty$

5. **Assumption FE5**:
   $Var(u_{it}|X_i, a_i) = Var(u_{it}) = \sigma_u^2$, for all t=1...T.

6. **Assumption FE6**:
   For all $t \neq s$, the idiosyncratic (unobserved) errors are uncorrelated (conditional on all explanatory variables and a_i):$Cov(u_{it}, u_{is}|X_i, a_i) = 0$.

Under the assumptions FE1-FE6 the fixed-effect estimator is BLUE (Best Linear Unbiased Estimator).

7. **Assumption FE7 - Normality**:
Conditional on X_i and a_i, the u_{it} are independent and identically distributed as Normal(0,σ^2).
Assumption FE7 assures us that fixed-effect estimator is normally distributed, its t and F statistics have exact t and F distributions respectively. Without the assumption FE7, we can rely on asymptotic approximation (however without further assumptions, they require large N - number of observations and small T - time periods). (Wooldridge (2015))

4.5 Optimizing the Model

Once the model is build for the first time, we have to further analyze its performance and assure ourselves, that all the assumption are met (Chapter 4.4). The most important would be to check the model for auto-correlation of any degree and check that the disturbances are homoscedastic. Each term is further explained in the subsections.

4.5.1 Auto-correlation

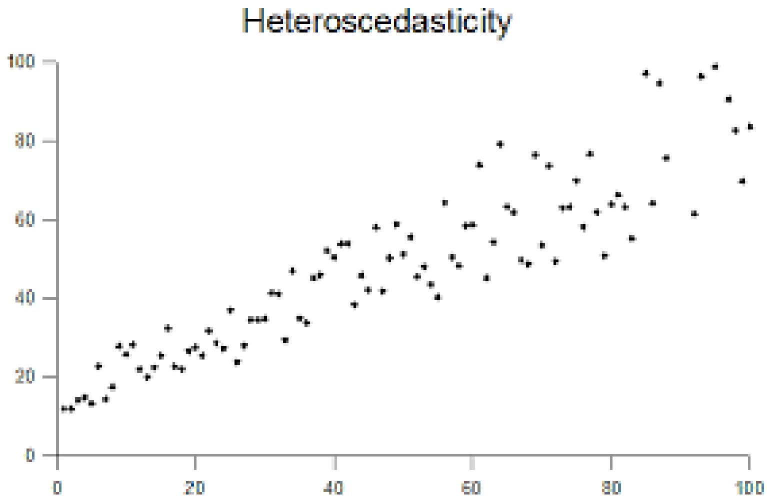
Auto-correlation or serial correlation is a phenomenon that can occur while working with time series. To define serial correlation we can use a definition: "Autocorrelation is a mathematical representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals" Wooldridge (2015). What it basically means is that there is a relationship between estimated processes from different time periods.
It can be measured by building a linear model, that plots residuals from different time periods against each other.The formula for this specific model is:
$residuals_{it} = \beta_0 + residuals_{it-1} + \eta_{it}$

If the model uncovers a statistically significant influence of lagged residuals on the non-lagged, we can say that the model suffers from serial correlation of the 1^{st} degree.

4.5.2 Heteroskedasticity

Heteroscedasticity "happens when the standard deviations of a predicted variable, monitored over different values of an independent variable or as related to prior time periods, are non-constant. With heteroskedasticity, the tell-tale sign upon visual inspection of the residual errors is that they will tend to fan out over time." (Engle (1982)).

Figure 4.1: Heteroscedastic Graph



Source: Engle (1982).

Chapter 5

Results and Discussion

5.1 Results Introduction

In this chapter we begin by studying descriptive statistics of our data. In the latter sections we move on to describing and interpreting our modelling results

5.1.1 Descriptive statistics

We will go trough all variables used in the analysis and discuss their meaning for the research. Please note, that brief description of the data and their sources are provided in the chapter 3. By each variable let us see its descriptive properties and then add our commentary.

5.1.2 Year

This variable has no specific interpretation value. It serves as a help variable to account for trend in other explanatory variables. Other meaning of this variable is to create unique indexes for working with panel data pooled OLS methods. In our formula notation, the Year variable stands for t.

5.1.3 Country

This variable is used to create panel data structure and enable us to work with panel data pooled OLS. In our formula notation, the Country variable stands for i. The list of countries presented in this paper can be found in Chapter 3.2 - Variable Desription.

5.1.4 Corporate Tax Revenue

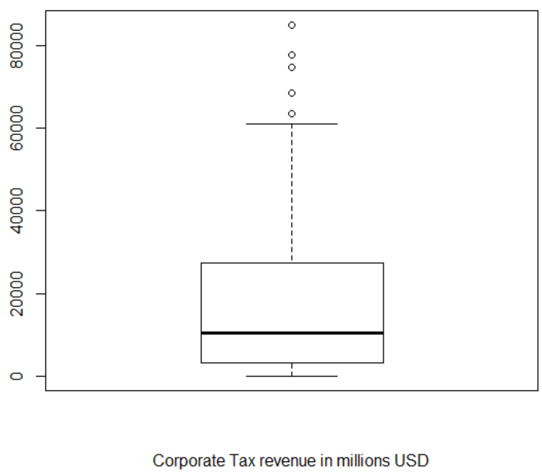
This variable represents total amount of tax revenue obtained in respective states.

Table 5.1

Mean	SD	1^{st} Qu.	Median	3^{rd} Qu.
18294	20396	3573	10555	27036

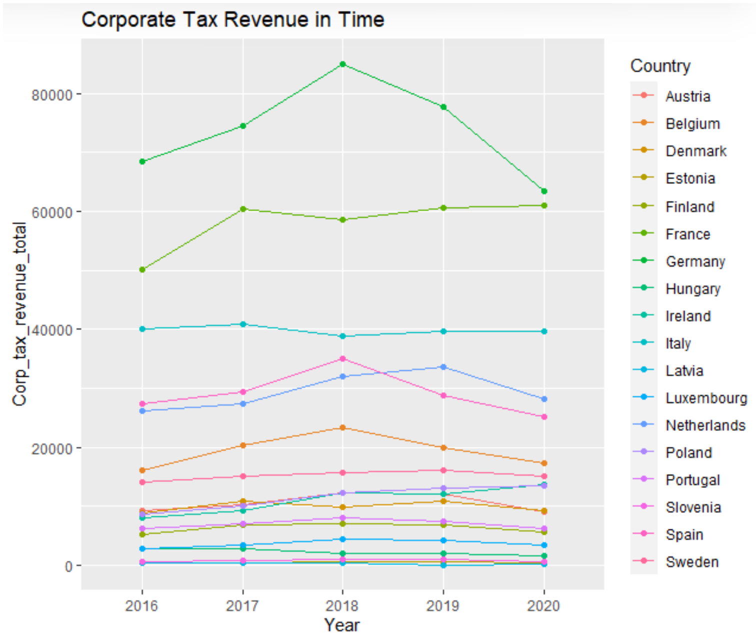
The average corporate tax revenue is approximately 18 billion USD and standard deviation is approximately 20 billion USD. This tells us that there are big differences between inspected economics or countries. However this is no surprise since the variable is measured in absolute numbers and there are rather large differences in size of the national economies as it is to be expected. Countries like Germany or France have disproportionately larger markets than for example Latvia or Estonia.

Figure 5.1: Corporate Tax Revenue - Boxplot



Upon inspecting distribution of the variable "Corporate tax revenue" we identified 5 outlier values. All come from Germany or France and since the variable is measured in absolute amounts the outlier is a result of robustness and size of the German economics. Therefore we decided that obtaining such values is reasonable and are not a product of error or unexpected disturbances.

Figure 5.2: Corporate Tax Revenue in Time



Even though it is safe to assume, considering GDP rises in time and population of European countries decline, so that GDP per capita will have an upward trend in time. However in a short time frame, there seems to be no visible trend and so there is no need for a correction or a detrending (transforming the data to get rid of the trend) .

5.1.5 GDP per capita

This variable is used to represent economic strength of the economies presented in the paper. We argue that GDP per capita can be translated to consumer strength and amount of their disposable income. Hence the "stronger" the market the more desirable and profitable it is for the MNEs to operate on such markets. Further more it is safe to assume that the economic performance of a country This variable represents total amount of tax revenue obtained in respective states. We expect, that this variable will have a positive influence on the dependent variable

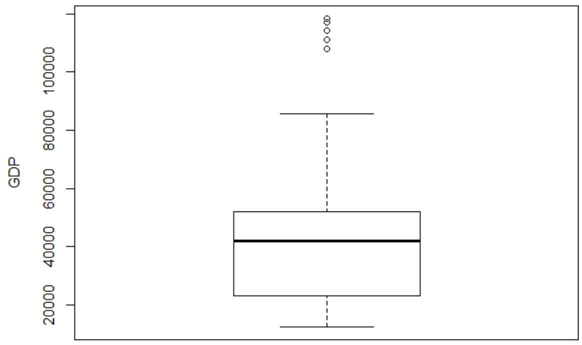
The average GDP per cap among countries represented in the study is approximately 42 thousand USD with a rather high standard deviation of 24

Table 5.2

Mean	SD	1^{st} Qu.	Median	3^{rd} Qu.
41973	24329.5	23173	41864	52038

thousand USD. Because this variables is calculated with respect to size of respective population, the total size of economy does not matter. However some countries in the data set are heavily oriented on specific industry or are officially home to unnaturally high amount of multinational enterprises due to benevolent tax code. Let us have a deeper look into the distribution of GDP per capita.

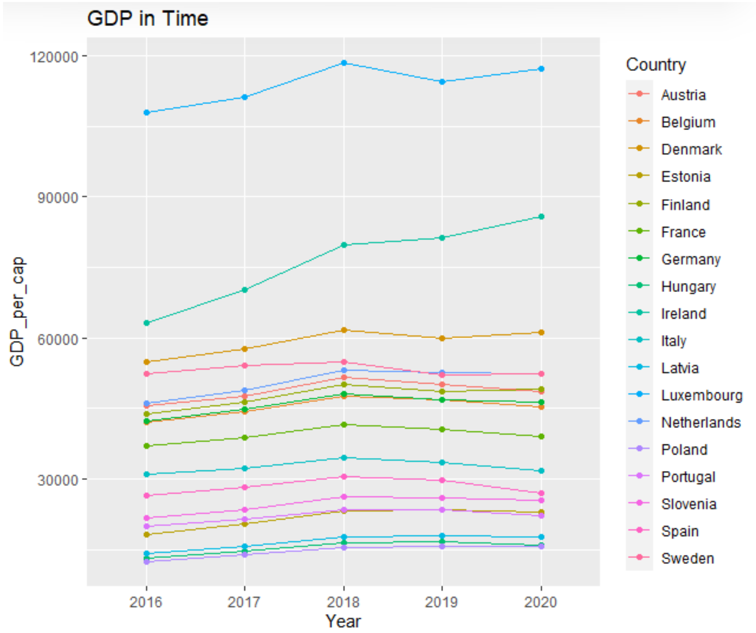
Figure 5.3: GDP per capita - Boxplot



From the box plot we can see that there are in total 5 outliers. All outliers belong to one state, that is Luxembourg. This can be caused by the fact that Luxembourg has its economy heavily oriented on banking sector and the massive GDP per cap is mainly caused by an influx of foreign investments. The Luxembourg financial sector was responsible for 86% of national GDP in 2014. (*The World Factbook (2014), Central Intelligence Agency*). Hence it is reasonable to assume that economy of Luxembourg is specifically influenced by its financial industry and the observations from Luxembourg were therefore excluded from further computations.

As discussed before, among the countries with highest value of GDP per capita is on one side Luxembourg and on the other hand are big and robust economies such as Germany or France. Among the lowest values are countries

Figure 5.4: GDP per capita in Time



such as Estonia, Hungary or Slovenia. There seems to be no visible trend in time in the studied period.

5.1.6 The Digital Economy and Society Index

This variable was chosen to be included in our model to reflect on digitalisation of the market, where multinational benefit from a digital infrastructure. We assume that the more digitalised a country is or the more robust digital infrastructure is present in a country, the more the digital companies benefit from it. Which would be reflected in our dependent variable. A high level of digitalisation on the other hand makes it also easier for digital MNEs to shift their profits to jurisdictions with lower effective or corporate tax rates. Either way expect the DESI to have a positive influence on the dependent variable. In context of our research this variable serves as indicator on how much of a presence do digital firms have in a respective country and what is the potential for such a service to exist in these regions. For methodology on how the DESI is derived, please consult with Chapter 3.

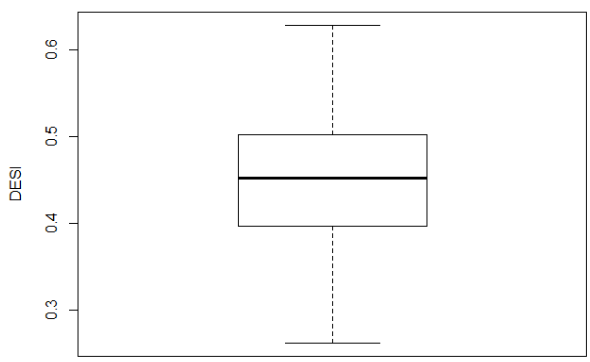
The average DESI in the selected European countries moves around 45%. This value is rather far above world average, amounts to 12% according to report of the European Comission (*European Commission, International Digital*

Table 5.3

Mean	SD	1^{st} Qu.	Median	3^{rd} Qu.
0.4512	0.08	0.3971	0.4516	0.5020

Economy and Society Index Final Report 2020). Among the most digitally advanced countries are northern countries like Finland, Denmark or Sweden. On the other hand Poland or Hungary are rank at the bottom of the list.

Figure 5.5: The Digital Economy and Society Index - Boxplot



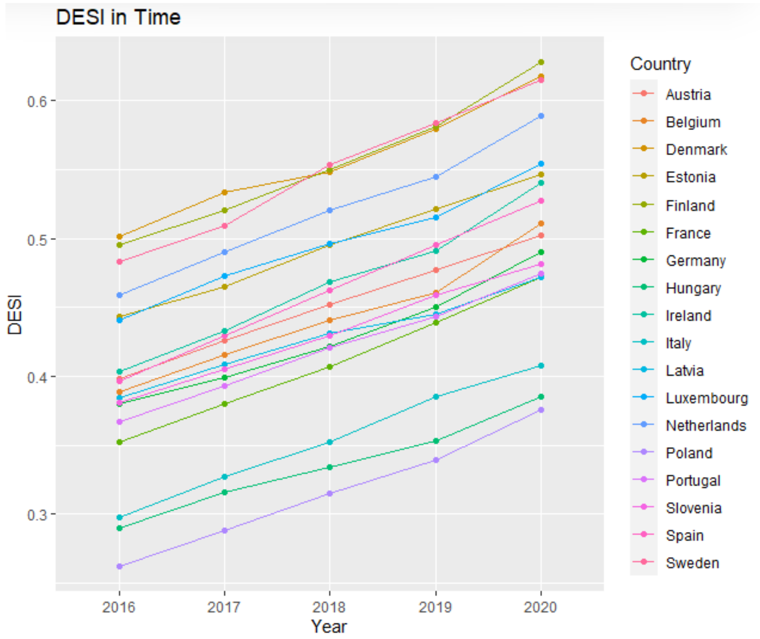
By looking at the box plot we can see, that there are no outliers. The variable is evenly distributed around its mean value.

Even though the variable has no outliers, after we plot the value of the index against time we can observe an obvious upward-oriented time trend in digitalisation of the European Union as can be expected (Figure). While building a model this trend has to be accounted for. Due to the trend a variable *Year* was added to the model.

5.1.7 International Tax Competitiveness Index

International Tax Competitiveness Index is a variable chosen for this paper for its unique interpretation value. While using this variable we think about multinational enterprises as a customers, who are looking for the "best deal"

Figure 5.6: The Digital Economy and Society Index in Time



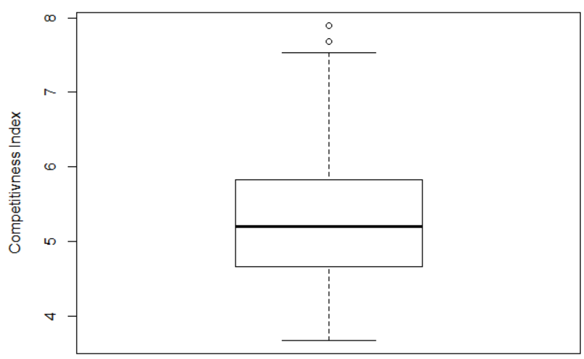
and the tax code or tax burden imposed on them by authorities adds to their costs. This variable therefore tackles our assumption, that some firms will be incentivised to shift their profits to regions with milder or "less expensive" tax codes (ie. less expensive in a sense, that tax is a cost for the MNEs). That means, that hypothetically the additional tax burden in form of digital service tax could hypothetically either lower the economic performance of a firm or make a profit shift more profitable for some digital firms. We expect that this variable will have positive influence on the dependent variable to some extent. For a certain level of International Tax Competitiveness index the influence is assumed to be negative. Let us have a look at the distribution of the variable.

Table 5.4

Mean	SD	1^{st} Qu.	Median	3^{rd} Qu.
5.281	0.95	4.682	5.281	5.818

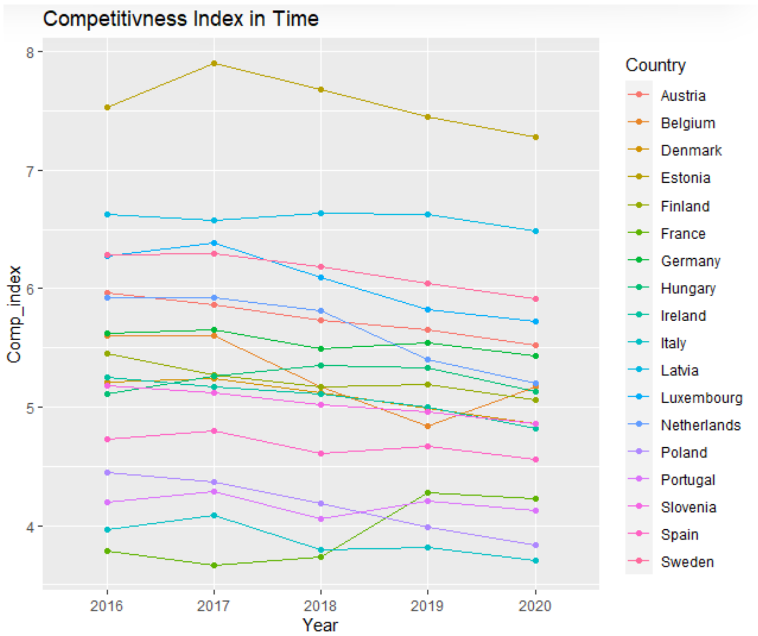
According to these descriptive statistics it is fair to say, that the total ITCI is mainly distributed among its average value. As mentioned before the most competitive tax code can be found in Estonia. On the other hand, countries like Italy, Poland or Portugal. For more detailed description regarding the methodology of this index, see Chapter 3.

Figure 5.7: International Tax Competitiveness Index - Boxplot



From the box plot we can identify two outliers, that have slightly higher value above the outlier border. Both values belong to Estonia. According to *Taxfoundation.org*, an organisation that calculates this index every year, Estonia belongs to some of the most tax-wise competitive economies in the world with its support for "free entrepreneurship and minimal bureaucracy" (see *BunnAssen for TaxFoundation.org, International Tax Competitiveness Index 2021*).

Figure 5.8: International Tax Competitiveness Index in Time



From the graph of the variable in time, we can see that there is a small down warding trend in the development of tax competitiveness. This seems reasonable since the countries compete between each other to influence multinational enterprises to set up headquarters or branches in the country and make it easier for these companies to operate across national borders.

5.1.8 is.DST

The *is.DST* is a dummy variable signalizing where and when an equivalent of a digital service tax was implemented. Among countries that were first to implement such a tax tool are France, Austria or Hungary. Unfortunately the tool is rather young and hence there is not a lot of observation that were influenced by an active digital service tax tool. The fact that there are all together only 8 observations create issues with statistical inference and reliability of the estimation.

5.2 Regression Model

In this section we are going to discuss properties of the fixed-effect model we chose to use as our best estimation method. Assuming all necessary assumptions are met (FE1-FE4, see Chapter 4 - Methodology), we can advance to discuss the results of our model and respective coefficients responding to predefined variables.

The final formula for the model is as follows:

$$log(Corporate\ Tax\ Revenue_{it}) = \beta_0 + \beta_1 * log(GDP\ per\ capita_{it}) + \beta_2 * DESI_{it} + \beta_3 * ITCI_{it} + \beta_4 * ITCI_{it-1} + \beta_5 * is.DTS_{it} + \beta_6 * Year_i + a_i + u_{it}$$

As a comparison, we also included first-differenced model to the analysis. The results of our estimation using fixed-effect and first-difference model are:

We retrospectively checked for auto-correlation by plotting residuals from different time periods against each other. We detected no auto-correlation of the model, since no residual estimators were statistically significant. To check for heteroskedasticity of the model we ran a series of Breuschâ€“Pagan tests. In both test we arrived at p-value of 0.17 and so we cannot reject the

Table 5.5: Regression Results

	Dependent variable:	
	Corporate Tax Revenue	
	FE	*FD*
	(1)	(2)
log(GDP.per.cap)	1.40143***	1.94647***
	(0.285)	(0.351)
DESI	-0.35223*	-6.12813*
	(2.661)	(3.367)
ITCI	-1.12706***	-0.93522***
	(0.218)	(0.194)
ITCI.lag	0.15633	0.10529
	(0.206)	(0.177)
is.DST	0.67941	0.81354
	(0.491)	(0.518)
Intercept	N/A	-0.10738
	N/A	(0.17151)
Observations	90	89
R^2	0.56653	0.6196
Adjusted R^2	0.51714	0.59522
Note:		*p<0.1; **p<0.05; ***p<0.01
FE = fixed-effect, FD = first-differenced		

null-hypothesis and we do not have enough statistical evidence to assume heteroskedasticity. After these tests and with believe that other assumptions are fulfilled, we can say, that the derived estimators are BLUE.

As expected variable *GDP per capita* has a positive and significant influence on the dependent variable. It makes sense to think, that economic performance of a country is influential with regard to amount of tax collected from such a economy. Due to the logarithmic transformation the β_1 coefficient can be interpreted as a percentage change (*Jeffrey M. Wooldridge (2015),Introductory Econometrics A Modern Approach, p.70*). In other words for every 1% change in *GDP per capita*, the dependent variable *Corporate Tax Revenue* increases for 1.4%(1.9% for FD model).
The Digital Economy and Society Index has a negative influence on the dependent variable. The coefficients (for both models) are significant at 10% significance level. The result is rather hard to precisely interpret, since the methodology used to compute the DESI is complicated and there are more factors, that could influence such variable. We discovered that the effect on the *Corporate Tax Revenue* is opposite to what we expected. It turned out, that the higher the digitalisation, the less tax revenue will be collected from the MNEs. As discussed above, a higher level of digitalisation can benefit digital corporate, but at the same time it makes profit shifting easier. Furthermore it is reasonable to think that due to higher level of digitalisation and internet access, the consumers in a country can access different markets more easily and look for "best offers" across borders. However analysis of such phenomenon is beyond the reach of this paper.
The International Tax Competitiveness Index turned out to also significantly influence the dependent variable (even at significance level lower than 1%). Surprisingly the effect in both models is negative in contrast to our initial hypothesis. We can argue that countries, that are trying to be among the most tax competitive ones, are also countries, that have rather smaller economies (such as Estonia) and they try to utilize this characteristic for their benefit. In comparison to big economies like Germany or France the influence of tax competitiveness is remarkably lower than influence of economic scale. To sum this interpretation up, smaller European countries tend to be more tax competitive than big European countries, however have far lower economic performance (measured by total amount of GDP). In hindsight, it would be probably better for the formula to chose different methodology to measure our dependent

variable in order to avoid such an ambiguous interpretation.
The lagged *The International Tax Competitiveness Index* turned out to be insignificant and so the hypothesis we hold turned out to be incorrect. One could argue, that the length of a period, which is necessary for digital firm to acclimatize to new tax code is perhaps longer than one year. This could be true and a subject of "How long does it take for firms to adjust to new tax code" could be further examined in future studies.
The intercept is missing in the fixed-effect model, due to the specific methodology of deriving estimators from this kind of a model. For first-difference model the intercept carries no interpretation value, since it is virtually impossible, that all the variables would be equal to zero.
Most importantly the variable *is.DST* came out of the model as insignificant at level of significance of 10%. However it is worth mentioning that if we were to expand the significance level to 15% or 20%, then the coefficient belonging to this variable would in fact turned out to be statistically significant. For both models the presence of DST positively influences the amount of *Corporate Tax Revenue* collected each year. If we were to approximate additional tax revenue using the fixed-effect model, the presence of a DST tool would raise the collected tax amount by 0.7%, which would roughly translate to total amount of $ 12 billion USD. This number is significantly different from US$ 1 trillion dollars per year, approximated by the European Commission (*European Commission (2020), Digital Taxation - state of play and way forward*). This significant difference can be caused by the fact that only 18 of 26 EU member states were included in this paper. Another possible explanation is that an implementation of a DST tool could have additional tax revenue effect, that were not covered by us. In addition a larger and more robust data set ranging along more time periods would be necessary if we were to arrive at conclusive results.
Finally we obtained a R-squared coefficient of 0.57 and 0.61 for our fixed-effect and first-difference model respectively. This is rather high R-squared given we try to estimate such a complex macroeconomic question as implementation of digital service tax, however in our case the R-squared coefficient is artificially enhanced by the methods we have utilised to derive the final regression estimators.

5.3 Shortcoming of the Analyses

It is fair to say, that there are some problems with the analysis that the reader should be careful about, when thinking about this study.

- The data set was rather small, counting only 90 observations. For example when using fixed-effect method, the model gets more efficient with increasing number of observations (the estimator is becoming more and more similar to true value of the coefficient). However it is questionable, whether 90 observations make the model efficient enough. We recommend to repeat this analysis in 2-4 year time with a larger data set. But it is hard to say whether the data set will in reality increase, because of upcoming global OECD/G20 initiative called BEPS, that should repeal digital service tax tool in participating jurisdictions with effect in 2023.
- In this work we use a generalised version of a digital service tax (DST). We should mention that the digital taxation rule differs in respective countries and tax jurisdiction. The digital taxation tool can have different tax rates or could affect slightly different parts of the economy and this paper does not account for such differences.

Chapter 6

Conclusion

To conclude our paper we would like to say that there are some hints that a significant relationship between presence of a digital service tax tool and amount of corporate revenue tax collected. The implementation of a digital service tax can positively raise the collected tax amount about 0.7%. This relationship is now according to the models only significant at level of 15%-20% and so to this fact and also due to small data sample, we would rather judge these results as rather inconclusive.

This study has also prepared a method to analyse this phenomenon in the future, when hopefully the data samples will be larger. Unfortunately there is also a chance that the data will be no longer collectable, since a Anti Base Erosion and Profit Shifting initiative from OECD/G20 will be implemented in 2023. Which means that the digital service tax tools will be repealed and there will be no data to be collected.

There is surely room for more research and finding ways on how to utilize such a special tool in order to combat challenges arising from global digitalisation of the economy. In our opinion, it would be interesting and very useful to deeply analyze behaviour of big digital companies on basis of their accounting details. This research could be used to better predict their behaviour and reaction to tax code changes and new implementations of tax laws.

Bibliography

Bánhidi, Z., I. Dobos, & A. Nemeslaki (2020): "What the overall digital economy and society index reveals: A statistical analysis of the desi eu28 dimensions." *Regional Statistics* **10(2)**: pp. 42–62.

Bunn, D. & E. Asen (2020): "International tax competitiveness index." *Dostupno na: https://taxfoundation. org/publications/international-tax-competitiveness-index* .

Burg, D. F. (2004): *A world history of tax rebellions: An encyclopedia of tax rebels, revolts, and riots from antiquity to the present.* Routledge.

COM, E. (2018): "Proposal for a council directive laying down rules relating to the corporate taxation of a significant digital presence. brussels, belgium. 21.3. 2018, 147 final 2018/0072 (cns)."

Corbo, V., P. Rojas *et al.* (1991): "World bank-supported adjustment programs: country performance and effectiveness." *Technical report*, The World Bank.

Data, O. (2018): "Oecd data." *Retrieved from OECD. org: https://data. oecd. org/tax/tax-revenue. htm* .

Dougherty, C. (2011): *Introduction to econometrics.* Oxford university press.

Engle, R. F. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation." *Econometrica: Journal of the econometric society* pp. 987–1007.

Jansky, P. *et al.* (2020): "Corporate effective tax rates for research and policy." *Technical report*, Charles University Prague, Faculty of Social Sciences, Institute of Economic â€¦.

Popescu, C. R. G. (2020): "Sustainability assessment: Does the oecd/g20 inclusive framework for beps (base erosion and profit shifting project) put an end to disputes over the recognition and measurement of intellectual capital?" *Sustainability* **12(23)**: p. 10004.

Russo, V. (2020): "Digital economy and society index (desi). european guidelines and empirical applications on the territory." In "Qualitative and Quantitative Models in Socio-Economic Systems and Social Work," pp. 427–442. Springer.

Wooldridge, J. M. (2015): *Introductory econometrics: A modern approach*. Cengage learning.

Appendix A

Title of Appendix A