Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

The Financial Secrecy Index:

An Information Theory Approach

Author: **Bc. Lukáš Galuszka**
Supervisor: **Petr Janský, Ph.D.**

Academic Year: **2015/2016**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently; using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 12, 2016

Signature

Acknowledgments

I would like to thank especially my thesis supervisor Petr Janský for his guidance, helpful comments and that he was always available when I ran into troubles. I would also like to thank the Financial Secrecy Index team of the Tax Justice Network, especially Moran Harari and Markus Meinzer, who were involved in the survey for this thesis. Lastly, I would also like to thank Nicholas Tyack for his advices that were very useful and helped putting the thesis together.

Abstract

The objective of this thesis is to evaluate alternative weighting systems to determine if they have the potential to improve the current weighting system of the Financial Secrecy Index (FSI). The FSI, a measure of countries' contributions to global financial secrecy, currently weights its 15 qualitative components equally. A web-based opinion survey conducted in January and February 2016 among academics, journalists, experts and other persons familiar with FSI serves as the baseline for assessing new weights. The new weights derived from the survey results are not significantly different from the equal weights in 14 out of 15 components. The survey results suggest that widely held opinion is consistent with equal weight assumptions. Statistical model selection criteria from information theory that penalize model complexity prefer in majority of cases the simple model over the more complex one even though more complex model provides better goodness-of-fit statistics. Alternative methods and analysis such as Principal Components Analysis is performed and discussed. The present work finds that, statistically, the weights should not diverge from the equal weighting system in use currently.

JEL Classification	F21, F36, F65, G15, G20
Keywords	financial secrecy, information theory, weight setting, composite index, opinion survey

Author's e-mail	galuszka.lukas@gmail.com
Supervisor's e-mail	jansky.peta@gmail.com

Abstrakt

Cílem této diplomové práce je analyzovat systémy určování vah indikátorů a zjistit, zda existují potencionální nové váhy jednotlivých komponentů v Financial Secrecy Indexu (FSI), které by vylepšily současně využívaný systém. FSI, jenž měří, jakým způsobem jurisdikce svou činností přispívají k světovému finančnímu utajení, nyní dává svým 15 kvalitativním komponentům stejnou váhu. Dotazníkový průzkum, který byl distribuován na internetu, a byl provedený v lednu a únoru roku 2016 mezi akademiky, experty, novináři a dalšími osobami obeznámenými s FSI, slouží jako základ pro určení nových vah. Nové váhy odvozené z výsledků průzkumu nebyly u 14 z 15 komponentů signifikantně odlišné od rovných vah. Tedy výsledky průzkumu ukázaly, že obecný názor mezi odbornou veřejností je konzistentní s používáním rovných vah. Statistická výběrová kritéria vycházející z information theory, která penalizují modely za jejich komplexnost, preferovaly ve většině případů jednodušší model nad složitějším modelem, navzdory faktu, že složitější model disponoval lepšími výsledky testů dobré shody. Tato práce také představuje alternativní metody, jako například analýzu hlavních komponent, jejichž výsledky interpretuje. Práce shledává, že váhy komponentů FSI by se na základě statistických výsledků neměly odchýlit od rovných vah, které se nyní užívají.

Klasifikace	F21, F36, F65, G15, G20
Klíčová slova	finanční tajemství, information theory, určování vah, složený index, průzkum názorů

E-mail autora	galuszka.lukas@gmail.com
E-mail vedoucího práce	jansky.peta@gmail.com

Contents

List of Tables

List of Figures

Acronyms

AIC	Akaike Information Criterion
BOPS	Balance of Payments Statistics
CDI	Commitment to Development Index
FSI	Financial Secrecy Index
GSW	Global Scale Weights
HDI	Human Development Index
ICOMP	Information-Theoretic Measure of Complexity
KFSI	Key Financial Secrecy Indicator
OFC	Offshore Financial Center
PCA	Principal Component Analysis
SIC	Schwarz Information Criterion
RMSD	Root Mean Squared Deviation
TJN	Tax Justice Network

Master's Thesis Proposal

Author:	**Bc. Lukáš Galuszka**	Supervisor:	Mgr. Petr Janský, M.Sc., Ph.D.
E-mail:	galuszka.lukas@gmail.com	E-mail:	jansky.peta@gmail.com
Phone:	774800331	Phone:	
Specialization:	FFM&B	Defense Planned:	June 2016

Proposed Topic:

The Financial Secrecy Index: An Information Theory Approach

Motivation:

Offshore financial centers, tax havens and financial secrecy in general are interesting phenomenons which effect the economies of the world at great scale as statistics suggest. Cross-border illicit financial flows amount to at least 1 trillion USD (World Bank, 2009) and in total there are more than 20 trillion US dollars located in secrecy jurisdictions (Tax Justice Network, 2012). It is essential to focus on the countries and jurisdictions that are major players of financial secrecy and the final destinations of these illicit capital flows. The depth of this problem is well depicted in the analysis for The Economist (Valencia, 2013).

The Financial Secrecy Index (FSI) focuses on ranking countries and jurisdictions according to their economic scale of provision of financial services (quantitative part) and also according to the policies, tax laws, international treaties etc. (qualitative part). These two dimensions together enable to rank small tax havens with very favourable tax environment and also large developed economies which fail in only few qualitative indicators.

In my thesis I will focus on the qualitative component of FSI. It assigns score for 15 Key Financial Secrecy Indicators (KFSI) that are weighted equally. There has been a debate whether the equal weights are the best fit (e.g. Smallwood, 2014). Deeper analysis of the weighting system will address this debate and create new weights that will better represent each component's importance. Information theory approach to these weights (similar research was conducted by Stapleton, Garrod, 2008) will add another perspective and help it more comprehensively shed light on the financial secrecy problem.

Hypotheses:

1. Hypothesis #1: Nominal (equal) weights do not match the real weights (coefficients) obtained from the data
2. Hypothesis #2: Equal weights do not provide best results, suggest the new weights of KFSI using opinion survey
3. Hypothesis #3: With respect to criteria of statistical model selection derived from Information theory that punish complexity, the newly established weights still provide better results

Methodology:

I plan to test the hypothesis #1 by comparing the nominal (equal) weights to the real ones that I will have obtained from results of regressions using the actual KFSI data. Multiple years will be taken into account as a robustness check. In order to test the second hypothesis I plan to carry out a survey that will be send to numerous researchers and academics and will assess the weights of each component within KFSI according to reserchers' opinions and subsequently compare them with equal weights. This survey will be constructed similarly to one conducted by Chowdhury, Squire (2006) – who proposed new weights that in fact provide better goodness of fit over the used equal ones for Commitment to Devepment Index (CDI) and Human Development Index (HDI) and whom I already contacted regarding their survey – and bearing in mind research about surveys published by Scott (1961). Adequate response rate in the survey will be also achieved by prearranged assistance from the FSI and focusing the survey on the right respondents.

I plan to determine whether the new weights are significantly different from the equal weigts and also assess the goodness of fit of suggested weights. Goodness of fit will be quantified using residual sum of squares and R^2.

Statistical model selection criteria rooted in Information Theory uses maximum likelihood function to determine goodness of fit. I plan to create such model and test hypothesis #3. Similar research was exercised by Stapleton, Garrod (2008) where interestingly, they found out that not equal weights suggested by Chowdhury, Squire (2006) for CDI which provide better goodness of fit are counteracted by the increased model complexity and they conclude that in overal equal weighting system for CDI should not be abandoned . I furthermore plan to compare the performance of this model with ones previously introduced in the thesis and ultimately decide the weights that are best fit for components of KFSI.

Expected Contribution:

My thesis will give new insights in financial secrecy and address the equal weighting system used in KFSI. New proposed weights as well as their analysis - if proven as a better fit for the FSI - could be applied to future versions of FSI that would enable producing more precise results of the indicators. Consequently, this could help with targeting financial secrecy problem in general more efficiently.

Outline:

1. Motivation: The financial secrecy in general creates an environment where huge amounts of capital are subject to no or very low taxes and furthermore enables illicit financial flows. In this section I will also discuss the Financial Secrecy Index and ways how to tackle this problem
2. Survey method and data collection: I plan to describe the details of conducting the survey and also the survey itself
3. Methodology: Theoretical background and methodologies that will be used to obtain results, also discussion and creation of models that will be used
4. Survey results: I will present the results from the survey and asses new weighting system
5. Results of the estimations: I will discuss the results of my analysis
6. Conclusion: Summary of the main findings of my thesis and their consequences for the FSI and discussion of the contribution my thesis will have brought

Core Bibliography:

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Proceedings of the Second International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 267–281.

Bozdogan, H., 2000. Akaike's information criterion and recent developments in information complexity. Journal of Mathematical Psychology 44, 62–91.

Chowdhury, S. and Squire, L., 2006. Setting weights for aggregate indices: An application to the commitment to development index and human development index. *Journal of Development Studies*, *42*(5), 761–771. doi:10.1080/00220380600741904

Cobham A., Janský P. and Meinzer M., 2015. The Financial Secrecy Index: Shedding New Light on the Geography of Secrecy. CGD Working Paper 404. Washington, DC: Center for Global Development. http://www.cgdev.org/publication/financial-secrecy-index-shedding-new-light-geography-secrecyworking-paper-404

Janský, P. 2013. Illicit Financial Flows and the 2013 Commitment to Development Index. CGD Policy Paper 034. Washington DC: Center for Global Development. http://www.cgdev.org/publication/illicit-financial-flows-and-2013-commitment-development-index

Lynn, P. and Jowell, R. (1996) How might opinion polls be improved? The case for probability sampling, Journal of the Royal Statistical Society. Series A (Statistics in Society), 159(1), pp. 21–8.

Murphy, R. 2008. Finding the screcy world: Rethinking the language of “offshore.” London: Tax Research LLP. http://www.taxresearch.org.uk/Documents/Finding.pdf ———. 2009. Where are the world’s secrecy jurisdictions? Tax Justice Network. Available online: http://www.secrecyjurisdictions.com/PDF/SJ_Mapping.pdf.

Myung, I.J., 2000. The importance of complexity in model selection. Journal of Mathematical Psychology 44, 190–204.

Scott, C., 1961. Research on mail surveys, Journal of the Royal Statistical Society. Series A (General), 124(2), pp. 143–205.

Smallwood, A. D., 2014. *An analysis of the efficacy of Tax Justice Network's methodology in constructing a secrecy index* (Report prepared for: Cayman Finance). Retrieved from http://www.caymanfinances.com/wp-content/uploads/2014/08/TJNreportFinal.pdf

Stapleton, L. M. and Garrod, G. D., 2008. The Commitment to Development Index: An Information Theory approach. *Ecological Economics*, *66*(2/3), 461–467. doi:10.1016/j.ecolecon.2007.10.009

Stapleton, L.M. and Garrod, G.D., 2007. Keeping things simple: why the Human Development Index should not diverge from its equal weights assumption. Social Indicators Research 84, 85–102.

Valencia, M. 2013. Storm survivors. Special reports of The Economist on offshore finance. The Economist, February 16. Available online: http://www.economist.com/news/special-report/ 21571549-offshore-financial-centres-have-taken-battering-recently-they-have-shownremarkable

Author **Supervisor**

1 Introduction

The world of financial secrecy first became prominent in the late 1960s and 1970s, when regulatory policies and attempts by advanced countries to control capital flows encouraged the shift of capital to banks and institutions in more friendly regulatory environments, into jurisdictions commonly known as ‘tax havens.‘ The financial services market grew tremendously in these tax havens from the 1970s on. This growth, combined with the low regulatory policies of these jurisdictions, has caused a rapid growth in the relevance of financial secrecy in the past half-century.

There is some evidence of the extent of the funds that flow through these “tax havens.” Cross-border illicit financial flows amount to at least 1 trillion USD (Baker, 2007) per year and in total there are more than 20 trillion US dollars located in secrecy jurisdictions (Henry, 2012). It is essential to focus on the countries and jurisdictions that are the major players of financial secrecy and the final destinations of these illicit capital flows in order to prevent them.

The Financial Secrecy Index (FSI) ranks countries and other jurisdictions by their secrecy and also by their share of the total exports of financial services in the world. Its primary use is to identify the largest players in the world of financial secrecy and illicit financial flows. It reveals that the traditional view of tax havens being small islands providing zero taxes for non-residents is misinterpreted. Large developed countries that are members of the OECD rank at the top of the FSI alongside a few of the more typical tax havens as we know them – small remote paradise islands. This identification of the true nature of the recipients and conduits of illicit financial flows is the first step and a foundation upon which real policies (e.g., sanctions) can be based in order to more effectively diminish the magnitude of illicit financial flows. This is a complex topic and for that reason the first chapter of this thesis is dedicated to explaining the world of financial secrecy and the FSI in general.

A comprehensive review of the literature is carried out to better understand the background of weight setting and constructing a survey including designing a questionnaire which is the crucial part that facilitates a good response rate of the survey. The methodology is presented at the end of this section.

In the next section, the statistical and analytical study of the dataset is presented. The dataset comprises of 102 jurisdictions and contains data regarding its scores from different FSI indicators and components from 2015. Descriptive analysis and principal component analysis (PCA) is carried out to allow for simplification of the index and to determine potential future directions for research.

The main reported FSI scores are calculated by combining quantitative and qualitative components. The qualitative component of the index is comprised of 15 parts that are each weighted equally. Equal weights are a simple solution to the weighting problem, but the approach does not necessarily provide the most accurate results. This thesis also aims at assessing a potential new weighting system by conducting an opinion survey to potentially replace the current equal weighting system. The proposed new weighting system is derived from the survey conducted as part of this study and is tested statistically.

The final chapter of the thesis applies model selection criteria from information theory to further specified models using different weighting systems (similar research was conducted by Stapleton and Garrod, 2008). The generated datasets are used to determine if the model using equal weights is favored over the models using unequal weights by goodness-of-fit statistics and model selection statistics from information theory. The thesis concludes with a discussion of the survey results and the testing of the different models, and emphasizes the implications that these results have not only for the FSI but for the issue of financial secrecy. Ultimately, it provides insight into possible future research.

2 The world of financial secrecy

The world of financial secrecy revealed itself during 1960s and 1970s when wealthy individuals and corporations discovered the possibilities of so called secrecy jurisdictions. Since then, the financial flows going through these jurisdictions grew enormously. Secrecy jurisdictions attract foreign capital by their anonymity and favorable tax rates and exempts. It is important to note that not all financial flows going through these offshore financial centers are illicit. However, the secretive financial environment enables fraud, tax-evasion and many other illegal financial activities. Wealthy individuals and corporations profit from these financial transactions dealt offshore as they do not have to disclose information, pay taxes and in general they are not obliged to follow strict rules that are enforced in the country of their business.

To put it into perspective, I let the numbers speak for themselves. For example the International Monetary Fund (IMF) (2000) indicated in their report that their calculations targeted at selected offshore financial centers on cross-border assets between offshore financial centers were about US$4,6 trillion at end-June 1999 that was about 50% of the world total cross-border assets. They specify that about US$1 trillion in located in the Caribbean and Asia.

Zoromé (2007) also adds some statistics in his paper. He suggested that offshore banks' assets amounted to US$1,9 trillion compared to US$16 trillion of total bank assets in 2003.These numbers are very likely underestimated as not all banks report to the Bank of International Settlements, which reported these numbers.

It is likely that almost US$21 trillion is dealt every year by wealthy individuals using the tax havens (Henry, 2012). Some reports go even deeper and suggest that the world's governments lose US$255 billion in revenues due to wealthy individuals depositing their capital and using the services of tax havens (Murphy and Vigueras, 2007). However, this number does not include the corporate funds. Governments lose what is

estimated around US$385 billion every year if the corporations are included in the calculations (Cobham, 2005). These numbers only reveal the losses for the governments. The total number of illicit financial flows in the world is estimated around US$1 – US$1,6 trillion (Baker, 2007).

Illicit financial flows, tax dodging and other related phenomena affect developing countries very dramatically. They are creditors to other countries and need to pay their obligations. It is estimated that revenue losses amounted up to US$50 billion for developing countries that were caused by the tax havens only in 2000 (Oxfam, 2000). Africa, the continent with probably the largest number of developing countries, lost around US$1 trillion in capital flight since 1970s when there was a boom of secrecy jurisdictions (Boyce and Ndikumana, 2012).

However, the tax evasion, frauds and corruption are also present in developed countries. For instance, in the past few years many European countries such as Portugal, Greece and Ireland are overwhelmed by their debt also due to their inability to collect taxes properly. In the previous paragraphs I briefly showed that addressing this issue is urgent and that it is a worldwide problem affecting all jurisdictions from poor third world countries to the wealthiest ones. In the following Chapter 2 the Financial Secrecy Index is discussed.

2.1 Financial Secrecy Index

The Financial Secrecy Index (FSI) is constructed every two years by team of individuals under the auspices of Tax Network Justice (TJN). TJN is an independent international network of individuals that specializes in research, analysis and advocacy of international tax laws and who are also interested in the area of financial regulation. The TJN team members study international treaties and tax laws and specialize in the effects that these policies have on the level of tax evasion and on illicit flows into tax havens.

The FSI ranks jurisdictions according to their level of financial activities on the global scale and also according to their level of financial secrecy. The level of jurisdictions’

financial activities on the global scale is determined by the Global Scale Weight (GSW) that for each jurisdiction represents a share of the provision of financial services that the jurisdiction has of the global total. Thus, it shows each jurisdiction's exports of financial services in comparison with other jurisdictions on the global scale in percentage. The part of the FSI mentioned second uses qualitative data about the jurisdictions to award them with a transparency score and subsequently computes the secrecy score for the jurisdictions. The more secretive a jurisdiction is, the higher secrecy score it gets. Both parts are discussed in detail further in this chapter.

The first FSI ever was published in 2009 and as already mentioned above, it has been constructed once every two years since (2011, 2013 and 2015). It is time demanding to construct a new FSI. Furthermore, all the data necessary for the computation of GSW (or construction of FSI in general) for a particular year are published by various agents with a delay (usually within year or two year after end of the particular year). Due to these facts the FSI for a particular year is constructed using the two-year old data. This means that FSI 2015 is constructed using the data for 2013, FSI 2013 is constructed using the data for 2011 and so on. This applies only for the computation and estimation of the quantitative part – GSW. It does not apply for the awarding of secrecy scores in the qualitative part. These scores are reassessed in the year in which the FSI is constructed. Consequently, the FSI 2015 is constructed with secrecy scores awarded in 2015 and GSW computed based on the data from 2013 – this provides a combination of up-to-date qualitative data and the completeness of quantitative data.

The FSI aims at identifying large players in the secrecy world. The FSI does not cover all the jurisdictions in the world but only the most important ones. This way the coverage is as large as possible while keeping in mind the constraints created by data and resource availability. The first FSI 2009 covered 60 jurisdictions that were picked from eleven listings issued by international bodies and academics. In FSI 2011 the aim was to include all jurisdictions that ranked among 20 jurisdictions with the highest GSW in 2009 (thus using two-year old data for year 2009 as explained above). Nine jurisdictions were added as a result of this criterion. The remaining four jurisdictions were added due to the fact that they were suspected to provide financial secrecy. In FSI

2013 there were seven additional jurisdictions added due to their position in the top 30 by Global Scale Weighting rank. In FSI 2013 there were two more jurisdictions included because of their high level of provision of financial secrecy. It amounts to a total of 82 jurisdictions in FSI 2013 (Financial Secrecy Index, 2015).

In the latest FSI 2015 the total number of jurisdictions increased to 102. Six jurisdictions were added because their exports of financial services were among the top 40 jurisdictions in the world. Another 7 jurisdictions were added as they have secrecy or tax haven ambitions. It was decided that all OECD members would be added to FSI 2015 and the remaining 7 that were not yet included entered the scope of FSI 2015 for that reason (Financial Secrecy Index, 2015).

Nevertheless, not including all jurisdictions means that qualitative secrecy scores are prepared only for the limited number of jurisdictions in the FSI scope. On the other hand, the GSW are always computed using the data for the whole world so the sum of GSW of jurisdictions is equal to 100%. But as previously mentioned – the aim is to cover the greatest number of jurisdictions that are the most important in terms of financial secrecy and level of their financial services exports. The 102 jurisdictions in FSI 2015 together comprise more than 98% of all global exports of financial services that is a primary data source when computing GSW (see later in this Chapter).

In general, these two parts combined provide a more complex insight into the world of financial secrecy. The qualitative indicators assess how much a jurisdiction is capable or willing to be an offshore financial center (OFC) or a tax haven. Furthermore, it can show that a jurisdiction enables illicit financial flows in general. The quantitative indicator shows what portion of the world's flows of financial services (exports of financial services) the jurisdiction mediates. The combination of the two different views on financial secrecy put into one final value of the index enables the comparison of large developed OECD countries that have rather low secrecy score with small so called tax havens - e.g islands with very small economies on the global scale but on the other hand these jurisdictions are very secretive.

The final FSI takes into computation the cubed value of Secrecy score and multiplies it with cubed root of GSW. Therefore:

$$FSI_i = SecrecyScore_i^3 \times \sqrt[3]{GSW_i} \quad (1)$$

The motivation behind this is to put more emphasis on the financial secrecy in the indicator. The FSI team provided another reasoning behind this - there is significantly more variation in the quantitative part (GSW) than in secrecy scores (Financial Secrecy Index, 2015). The final ranking of the FSI 2015 is appended in Appendix 1.

Cobham, Janský and Meinzer (2015) study the top ten jurisdictions ranked by FSI, Secrecy Scores and GSW. The table presented in the paper contained data for FSI 2013 but the following Table 1 shows the top ten rankings for FSI 2015.

Table 1: Top ten jurisdictions by FSI, Secrecy Score & GSW

Ranking by	FSI	Secrecy Score	GSW
1	Switzerland	Vanuatu	USA
2	Hong Kong	Samoa	United Kingdom
3	USA	St Lucia	Luxembourg
4	Singapore	Liberia	Germany
5	Cayman Islands	Brunei Darussalam	Switzerland
6	Luxembourg	Antigua & Barbuda	Cayman Islands
7	Lebanon	Marshall Islands	Singapore
8	Germany	Bahamas	Hong Kong
9	Bahrain	Nauru	France
10	UAE (Dubai)	Belize	Ireland
Avg Secrecy Score	68	81,8	57,4
Sum of GSW	56,48	0,079	78,63

Source: Financial Secrecy Index 2015

Even when secrecy scores are cubed so there is more emphasis on the secrecy part, there is no jurisdiction from the top ten jurisdictions ranked by secrecy scores that would rank in top ten jurisdictions ranked by the overall FSI. Furthermore, the top ten jurisdictions ranked by secrecy scores together comprise of less than 0,1% of the total exports of financial services. This demonstrates the importance of combining two parts of the FSI. The following two sections are dedicated to both thhe quantitative and qualitative parts in detail.

2.2 The Global Scale Weights (quantitative part)

In this chapter I briefly describe the quantitative part of FSI – the computation of GSW. It is computed for each jurisdiction as a share of its exports of financial services to the global total.

$$GSW_i = \frac{exports\ of\ financial\ services_i}{sum\ of\ all\ jurisdictions' exports\ of\ fin.services} \quad (2)$$

These global weights are based on publicly available data and the missing data are extrapolated. This idea originated in Zoromé (2007) where he first came up with using exports of financial services to identify offshore financial centers. He also pioneered the extrapolation of the missing data for certain jurisdictions. It is especially the jurisdictions that harbor illicit financial flows and provide financial secrecy that do not declare their statistics. Therefore, their data are missing the databases.

The primary source of data is Balance Of Payments Statistics (BOPS) that are extrapolated according to Zoromé (2007). These true non-adjusted data cover around 50% of the sample of all jurisdictions (247 jurisdictions in 2015).The second and third preferred data sources are the extrapolated asset on stocks of portfolio data from International Investment Position (IIP) and the Coordinated Portfolio Investment Survey (CPIS). Both IIP and CPIS are published by the International Monetary Fund (IMF). The fourth form of extrapolation uses the CPIS liability data due to the lack of asset data availability for extrapolation. Unfortunately, data for almost 30% of jurisdictions are obtained by this extrapolation which provide probably the least accurate results of all data sources.

The GSW itself does not serve as an indicator of any financial secrecy or the magnitude of illicit flows. It only geographically determines where most of the financial services flows originate. Only when combined with the qualitative indicator of secrecy scores to create the FSI, it shows the distribution of illicit flows and financial secrecy in the world.

The main insufficiency of this methodology is the lack of data available in the publicly available databases. The extrapolation is a means of obtaining some kind of data for the majority of jurisdictions. However, extrapolated data of course cannot be a match for true data but they provide completeness and robustness of the dataset. The GSW are not the subject of this thesis and will not be discussed further. For more details refer to Financial Secrecy Index (2015).

2.3 Secrecy scores (qualitative part)

As already previously mentioned, the so called secrecy scores are awarded to each jurisdiction based on 15 Key Financial Secrecy Indicators (KFSI) that are discussed further. These KFSI are based on a fully referenced publicly available dataset created by FSI on www.financialsecrecyindex.com. All indicators were compiled based on the score awarded according to the qualitative data contained in the documents, laws, regulations, international treaties and many more reports etc. Each indicator consists of sub-indicators that each has different effect on the final score. Some of them are based on a simple YES/NO answer but some are more complex and the awarding of the score is individual. This thesis does not go into great detail about the sub-indicators as it is a very comprehensive subject that could be the topic of another thesis. However, this thesis focuses on setting weights for each of fifteen KFSI and therefore each indicator is briefly analyzed (Financial Secrecy Index, 2015).

Fifteen KFSI can be divided into 4 groups according to the dimension of the secrecy:

- Knowledge of beneficial ownership (3 indicators)
- Corporate transparency (3 indicators)
- Efficiency of tax and financial regulation (4 indicators)
- International standards and regulation (5 indicators)

Secrecy scores are awarded via transparency credits. A jurisdiction can obtain a minimum of 0 and a maximum of 1 transparency credit from each of the 15 KFSIs.

These KFSIs are all equally weighted. Put differently, each KFSI has its weight equal to 1. Fittingly, if a jurisdiction is transparent according to the criteria of indicators then it is awarded transparency credits. Each jurisdiction can earn an equal number of transparency credits making their comparison straightforward. Nevertheless, as previously mentioned, the final values from which the FSI is constructed are not transparency credits but secrecy scores – an opposite of the transparency credits expressed as a percentage. It is computed as follows:

$$SecrecyScore_i = \frac{15 - \sum transparency\_credits_i}{15} \times 100\% \quad (3)$$

The term secrecy credits will also be used further in the thesis. Secrecy credits are the opposite of transparency credits and are computed as 1 – transparency credit. A jurisdiction that is very secretive will not be awarded any transparency credits but its secrecy credits will be high. On the other hand a jurisdiction that has the maximum transparency points (100%) will have 0% secrecy credits and also a secrecy score of zero. Appendix 2 lists all fifteen KFSI grouped into the listed dimensions of secrecy that are listed above with short descriptions.

3 Setting of the weights

As previously mentioned, each of the fifteen KFSI has the same weight equal to one. This means that each indicator influences the final secrecy score for the jurisdiction by the same scale. The current equal weighting system is subject to criticism and much discussion. Generally, the equal weights are considered a simple solution and are widely used for their simplicity and defended because all components are equally important (Decancq and Lugo, 2010). Sometimes, it would be ideal to abandon the equal weight assumption but there is no basis for creating different weights (Mayer and Jencks, 1989). Finding the basis is not a simple task but there are various methods of setting weights discussed further in this chapter.

On the other hand, there are hardships when using equal weights. Chowdhury and Squire (2006) summarize the ideal approach as one that weighs each component by its impact on the ultimate objective. There are not only general statements about the use of equal weights but there are also concrete comments on the weights used by the FSI. Smallwood (2014) directly comments on the KFSI weighting system and believes that it is not likely that each component has the same weight.

Nevertheless, the choice of weights always has to be justified and a scientific basis or reasoning provided. It also should be transparent and open to discussion and questioning in public (Anand and Sen, 1997).

This thesis revisits the current equal weights and ultimately aims to suggest a better alternative to the current weighting system. The following paragraphs review different methods of creating weights and different weighting systems. The literature is reviewed and the examples of the application of each weighting system are discussed as well.

3.1 Approaches to weight setting

There are many different approaches to setting of the weights. Decancq and Lugo (2008), Decancq and Lugo (2010) survey different weighting methods and systems used in the various multidimensional indices of well-being. Even though they solely focus on various well-being indicators, they introduce three classes of approaches to set weights that can be used universally:

1. Normative
2. Data Driven
3. Hybrid

Normative approaches do not set weights according to any statistics but rather the weights are set based on a judgment. They set weights in accordance with what it is believed that the weights should be to best fit the needs of the indicator. The problem is that the acumen made about the weights is often individual and does not have to represent the reader's opinion. This drawback is identified as paternalism. As already claimed in the impossibility theorem (Arrow, 1950), Decancq and Lugo (2010) follow by stating that it is not possible to aggregate individuals' opinions into one result. While normative approach to setting weights puts judgments and one's opinions as a baseline for setting weights, the second class of weighting approaches focuses on the descriptive side of the data and analysis – the data driven approaches.

Data driven (or descriptive) weights are not based on judgment or opinion. They are either a function of some distribution or e.g. can be based on the real effects of each indicator within multidimensional indices. Descriptive approaches sometimes use econometric regressions and in general set the weights according to the data or statistics – thus they set the weights based on statistics and omit individual opinions or judgment.

There is usually a significant difference between normative and data driven (descriptive) weights. The problem that arises from this difference is called Hume's law (or Hume's guillotine) (Hume 1740). This issue was already discussed in 18th

century and similar ideas were elaborated e.g. by Decancq and Lugo (2010). Simply put, the data-driven approach weighs components by what "is" and normative weights components by what "should be". Hume's law has been revisited multiple times by numerous researchers and academics and it has been also formulated by Popper (1948) and others.

The third class of hybrid approaches combines aspects of both of the mentioned approaches. These approaches aim to eliminate the drawbacks of both normative and data driven setting of the weights. The disadvantage of these hybrid approaches is usually the great difficulty of the setting. Sometimes, instead of eliminating the drawbacks of each approach, they suffer from both types of them. They usually follow certain distribution of the data and combine it with individuals' opinions. The hybrid approach will be further discussed in the following chapters as well as the first two mentioned approaches.

3.2 Normative approaches

Normative approaches rely on the opinions of individuals and their judgment rather than the actual data. Equal weights, opinion weights and price-based weights – these are three different approaches to setting weights that are all considered to be normative.

Equal weights are in general used most frequently and are very common. The reason is that they are quite simple. The equal weights have already been discussed at the beginning of this Chapter. The equal weights have been mentioned and/or discussed in Justino (2005), Lugo (2007) and also in Maasoumi and Lugo (2008).

The weights can be based also on opinions of experts from the field or individuals that are interested in the topic. These approaches adjust weights to the format how they "should be" according to opinions or economic intuition. These weights also represent a clear opinion of the scientific community (Mascherini and Hoskins, 2008).

There are two types of creation process of the weights according to expert opinions: the Analytic Hierarchy Process and the Budget Allocation method. The Analytic Hierarchy Process was pioneered by Saaty (1987). This method has been also a subject

of Rabbani and Rabbani (1996) and also Hummel (2001). In general this method asks questions “Which indicators are important and how much?” (Decancq and Lugo, 2010). The results are put into the matrix from which the results are derived. An example of these weights can be Nardo et al.(2005) where the authors focus on the composite indicators.

The second method is Budget Allocation. Each participant distributes certain number of points to each indicator that are then transformed into weights. Chowdhury and Squire (2006) used this method to set new weights for development indicators Health Development Index (HDI) and Commitment to Development Index (CDI). They chose the development researchers and academics from the whole world as the expert population. Similar research was conducted by Moldan and Billharz (1997) that surveyed 400 German experts on their opinion about the indicators that contribute to air pollution. Another example of Budget allocation method when setting opinion based weights is presented by Mascherini and Hoskins (2008).

Lastly, the price-based weights also provide a method of setting new weights. However, this method is used rarely and it is argued that it is not appropriate for multivariate indices (Sen, 1997). Price based weights are derived from marginal rates of substitution. Even though this method is not very popular for multivariate indices, there are exceptions. Becker et al (2003) used it to compute income equivalent compensations. Fleurbaey and Gaulier (2009) who studied living standard, incorporated this method to assess weights to unemployment, life expectancy and others.

3.3 Data driven approaches

Weight setting that uses statistical methods, analyzes frequency or derives weights from description of dataset belongs to a category of descriptive approaches (which is synonymous to data driven approaches).

Probably the most common statistical approach used to describe the data is the one using Principal Component Analysis (PCA). In short, PCA transforms the initial

variables into equal number of new variables that are not correlated and it is easy to determine what proportion of variance new variables explain. The weights are determined using linear combinations (Decancq and Lugo, 2010). This method was used e.g. by Klasen (2000) on measuring poverty and deprivation in South Africa. This method is suitable for the models where variables are highly correlated and a new set of variables created using PCA removes the correlation since the variables are not correlated.

Respondents sometimes criticize the weighting method picked by the researchers. This problem is addressed by most favorable weights that allow each respondent to create his own weighting system. This way each respondent is comfortable with his responses but on the other way it is very difficult to compare these results and generalize the outputs into comprehensible results.

Other data driven methods include those using a frequency of the data using a function of the distribution (applied e.g. by Brandolini, 2008 or Desai and Shah, 1988). Factor analysis, a simple approach that awards weights based on the observed data and descriptive statistics also belongs to this category of approaches and was used by Fusco and Dickes (2008) and others.

In general, the descriptive weights provide a useful tool for generating weighting schemes based on the dataset utilized. In certain cases they can provide suitable results but they overall suffer from lots of drawbacks and limitations.

3.4 Hybrid approaches

Hybrid approaches offer either a combination of the two general methods of weight setting mentioned in the previous paragraphs or generate weights by a method that did not fit into previous categories. They can benefit from the advantages of combining few approached together but they can also suffer from the drawbacks of different approaches. One must carefully analyze hybrid weight setting in order to prevent the second mentioned possibility. There are two basic types of hybrid weights: Stated preference weights and Hedonic weights.

The stated preference weights are somewhat similar to opinion based weights. The respondents are usually asked to assign ranks to variables and the weights are computed via a function. Additional information is collected with the ranks which allows the practitioner to avoid Hume's guillotine (Decancq and Lugo, 2010) that was discussed at the beginning of this Chapter 3. Stated preferences are frequently used in environmental economics (Carson and Louviere, 2011). They were used by De Kruijk and Rutten (2007) to construct a composite poverty index that was weighted based on stated preferences, and many others.

The hedonic method uses regression to set appropriate weights. It is necessary to construct a model where the dependent variable is tested against the variables that can be described as its characteristics. Various regression methods allow controlling and cleaning the results but on the other hand it can suffer from problems including large standard errors, collinearity and the issue of choosing the right variables to use based on their significance.

As already mentioned in this chapter, ultimately, the normative approach has been chosen. The opinion survey was conducted similarly to Chowdhury and Squire (2006) and is discussed in detail in corresponding chapters. The normative approach is preferred over the data driven one because we do want to keep the fifteen KFSIs and since they are qualitative indicators, the weights distributed according to experts' and academics' opinions best represents how the weights in fact "should be."

Furthermore, the analysis using information theory that penalizes complexity of the models is prepared in Chapter 6.4. Stapleton and Garrod (2007) and Stapleton and Garrod (2008) conducted a similar analysis to identify new suggested weights proposed by Chowdhury and Squire (2006) for CDI and HDI respectively.

4 Description of dataset and analysis

This thesis focuses on weighting fifteen KFSI and thus the dataset for descriptive analysis is quite simple. The weights that will be assessed are normative and not data driven, thus there is no need for the detailed data analysis for purposes of setting the weights. However, there has been a criticism formulated by Smallwood (2014) that criticizes the FSI for not including any analysis of the components or indicators.

As mentioned in the previous chapters, The FSI is constructed every two years. So far, the number of jurisdictions covered and ranked by FSI increased with each new published FSI. In 2009, the FSI covered 60 jurisdictions, in 2011 it covered 73 jurisdictions. In 2013, the coverage increased to 82 and the latest FSI 2015 ranks 102 jurisdictions. Furthermore, in 2011, the FSI significantly changed the methodology for awarding secrecy scores and KFSI. Additionally, with every new FSI published, the KFSIs are revisited and minor changes are made. From these facts arose a problem – how to select the right dataset.

Two possible approaches have been identified. First is to use only the latest data on 102 jurisdictions from 2015. These data provide current information on the largest number of jurisdiction possible from the FSI history and benefit from years of experience while determining the values of various KFSI. Unfortunately, this dataset only provides a cross-sectional analysis – an analysis at one specific time. On the other hand, the second approach is to create a dataset using panel data on only 60 jurisdictions. This dataset benefits from using multiple years but suffers from inconsistency that is caused by minor changes to KFSIs and awarding secrecy scores throughout the years. Due to the fact that the methodology is adjusted with every new version of FSI, the cross-sectional data from 2015 is chosen. This is in line with Stapleton and Garrod (2007) and Stapleton and Garrod (2008) that also use cross-sectional data for their information theory approaches to new weights set for CDI and HDI by Chowdhury and Squire (2006).

The dataset contains secrecy credits for fifteen KFSIs for each of 102 jurisdictions and also final FSI value for each of them except for one – Nauru (the data could not be extrapolated due to highly opaque economy and thus the GSW and consequently the FSI could not be prepared). This dataset does not include the sub-indicators of KFSIs as this thesis focuses only on aggregate KFSI and not their details. It is essential to note that the FSI 2015 results do not provide results of secrecy score for 9 jurisdictions (see Appendix 1). The FSI 2015 index was assembled during 2015 and not all necessary data and resources were available. However, this thesis uses more recent data where secrecy scores are available for all 102 jurisdictions presented in FSI 2015. The ranking of the countries and their overall FSI is thus different from the FSI 2015 results. However, this does not affect the thesis results at all as the FSI results are used only as a reference numbers. Still, the more recent data are preferred as they reflect the latest changes in the world and provide more extensive dataset of 102 jurisdictions as a bonus.

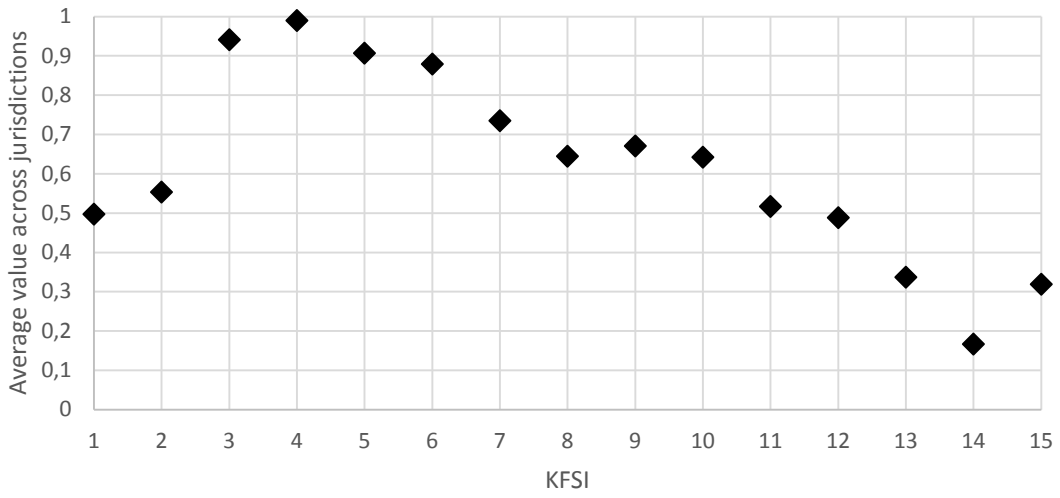
The descriptive analysis is done, simple regression is run and Principal Components Analysis is performed.

4.1 Descriptive statistics

The descriptive analysis is performed in order to determine any patterns or useful insights that can be used later in assessing weights. The Figure 1 represents the averages of secrecy scores awarded to jurisdictions for all KFSIs. Jurisdictions were the most transparent for KFSI14 (International Transparency Commitments). This suggests that most of the jurisdiction ratified the five most relevant international treaties relating to financial transparency. On the other hand, almost all jurisdictions scored full secrecy point for KFSI3 (Recorded Company Ownership) and KFSI4 (Public Company Ownership). In other words almost all jurisdictions do not obtain and keep updated details of the beneficial ownership and ownership of public companies available online or at reasonable cost. Thus these two components serve as a placeholders and not have any real impact on the ranking of the countries.

Source: Financial Secrecy Index 2015

Figure 1: Average values of secrecy scores for each KFSIs across jurisdictions



Source: Financial Secrecy Index 2015

4.2 Regression analysis

A simple OLS regression is run to show significance of the each variable (KFSI) and how much it is correlated with the final FSI value for the jurisdiction. There are hybrid approaches that use the regression results to assess the weights for components.

$$FSI_i = \beta_0 + \beta_1 kfsi1_i + \beta_2 kfsi2_i + \beta_3 kfsi3_i + \beta_4 kfsi4_i + \beta_5 kfsi5_i + \beta_6 kfsi6_i + \beta_7 kfsi7_i + \beta_8 kfsi8_i + \beta_9 kfsi9_i + \beta_{10} kfsi10_i + \beta_{11} kfsi11_i + \beta_{12} kfsi12_i + \beta_{13} kfsi13_i + \beta_{14} kfsi14_i + \beta_{15} kfsi15_i \quad (4)$$

The dataset contains transparency credits for each jurisdiction awarded for each KFSI. This means that all coefficients should be intuitively negative. If the values of transparency credit are to be transformed to show secrecy instead of the transparency, it could be done simply replacing values with “1 – transparency credit.” Then the values of coefficients would not change only they would be all positive instead of negative. The final result remains unchanged and therefore this procedure will not be done. Simple OLS is run with transparency credits as independent variables. Negative coefficients are expected. Table 2 shows the output from the statistical software Stata.

Table 2: Results of OLS regression

Independent variables	Coefficient estimate	SE	T
constant	55,13	249,96	0,22
KFSI1	-340,16	269,55	-1,26
KFSI2	-137,66	132,95	-1,04
KFSI3	-2953,78	2131,91	-1,39
KFSI4	-671,07	994,01	-0,67
KFSI5	-41,27	146,75	-0,28
KFSI6	1379,11	1061,45	1,30
KFSI7	-72,62	102,54	-0,70
KFSI8	-21,06	119,58	-0,18
KFSI9	-35,19	104,49	-0,34
KFSI10	73,09	162,35	0,45
KFSI11	850,98*	290,15	2,93
KFSI12	-170,93	103,46	-1,65
KFSI13	134,78	131,41	1,03
KFSI14	77,71	273,86	0,28
KFSI15	-55,06	222,72	-0,25
	R-squared = 0,2605		

* Significance of standardized coefficients at the 95 % confidence level
Source: own estimation; Stata software output

The R^2 of the model is very low which indicates the model is not well constructed. Only KFSI 11 is significant and all other KFSIs are not significant. Additionally, I expected all coefficients to be negative. Instead, there are 5 KFSIs with positive coefficients. We can conclude that the simple OLS regression is not a right tool for this kind of analysis and further different analysis is discussed later (different methods and different dataset). On the other hand this regression hinted that each individual KFSI have little impact on the final FSI value. The final secrecy score is cubed when the final FSI value is computed which could also be the reason why this model did not provide good results. This also implies that regression based weights would not be the right tool for this index.

4.3 Principal Components Analysis

There are fifteen KFSIs which is a large number for a detailed analysis. Furthermore, there can be bias in datasets with high number of variables. Also, multicollinearity can be present and some variables might be even insignificant. The Principal Component Analysis (PCA) is a method for simplifying datasets. It was pioneered by Pearson (1901) and then significantly revisited by Hotelling (1933).

PCA linearly transforms current variables into the new set of variables (the principal components) that are linear function of the current variables and are mutually uncorrelated (Sewell, 2008). The new variables are called “principal components” and only the most important ones are used to estimate the new simpler model that simplifies the analysis. Also, PCA can be used to set new weights of the components as previously discussed in the previous Chapter.

The PCA was done in software Gretl by estimating a covariance matrix and then performing the PCA. The following Table 3 represents the output of PCA:

Table 3: Eigenanalysis of PCA

Component	Eigenvalue	Proportion	Cumulative
1	5,5486	0,3699	0,3699
2	2,0605	0,1374	0,5073
3	1,5984	0,1066	0,6138
4	1,1321	0,0755	0,6893
5	0,8934	0,0596	0,7489
6	0,7641	0,0509	0,7998
7	0,5952	0,0397	0,8395
8	0,5359	0,0357	0,8752
9	0,4585	0,0306	0,9058
10	0,4032	0,0269	0,9327
11	0,3379	0,0225	0,9552
12	0,2524	0,0168	0,972
13	0,2333	0,0156	0,9876
14	0,1791	0,0119	0,9995
15	0,0074	0,0005	1

Source: own estimation; Gretl software output

If the purpose of this analysis was to assess new weights – one could further derive

them from the 15 new principal components that were created in this process. The components with Eigenvalues[1] larger than one are further analyzed. I identified four of these components. The first four principal components combined explain almost 69% of the variance while the first principal component explains 37% of the variance. The following Table 4 depicts how the selected four principal components are correlated with the original variables – KFSIs.

Table 4: Eigenvectors of the PCA (correlation of PC with KFSIs)

	PC1	**PC2**	**PC3**	**PC4**
kfsi1	0,266	0,007	0,329	0,128
kfsi2	0,127	-0,352	-0,078	-0,389
kfsi3	0,338	0,169	-0,371	-0,029
kfsi4	0,081	0,052	0,412	0,30
kfsi5	0,272	0,182	-0,348	-0,001
kfsi6	0,342	0,160	-0,364	-0,032
kfsi7	0,244	-0,381	0,191	0,184
kfsi8	0,328	-0,275	0,128	0,017
kfsi9	0,237	-0,351	0,092	0,020
kfsi10	0,186	-0,424	-0,127	-0,267
kfsi11	0,201	0,314	0,330	-0,468
kfsi12	0,284	0,253	-0,003	0,282
kfsi13	0,296	0,162	0,156	0,255
kfsi14	0,316	-0,035	-0,086	0,127
kfsi15	0,189	0,269	0,327	-0,506

Source: own estimation; Gretl software output

Each of four PC is correlated with different KFSIs. The name and characteristic of each PC is determined from set of original variables that are highly correlated with given PC. It is observable that PC1 is highly correlated with following KFSIs: 3 (Recorded Company Ownership), 6 (Country-by-country reporting), 8 (Efficiency of Tax Administration) and also 14 (International Transparency Commitments). It is difficult to see what these indicators could have in common as PC1 is highly correlated with one indicator from each class previously presented. However, if one was to look at the

[1] For details about eigenvalues and eigenvectors refer to Hazewinkel, Michiel, ed. (2001)

dataset and points awarded for these KFSI, it unveils the common denominator: all jurisdictions are among the wealthier jurisdictions and are economically well developed. Thus the appropriate name could be: “Jurisdiction is economically developed.” The second PC is highly correlated with KFSIs number 7, 9 and 10 (Fit for information exchange, Avoids Promoting Tax Evasion, Harmful Legal Vehicles). All of these KFSI are grouped into the “Efficiency of Tax Administration” that in turn could be appropriate name for the new variable.

Principal component number three is correlated with KFSIs 4, 5 and 6 which coincidentally are three indicators in the “Key aspects of corporate transparency regulation” class. Thus the name for this variable could be “Good corporate transparency regulation.” This name was derived from the correspondent class labelling. The last of the new variables (PC4) is highly correlated with KFSI 11 and 15 and also with other KFSIs from the “International standards and cooperation” class. That is why the author suggests naming this new variable “Jurisdiction cooperates on the international level.”

The PCA allowed us to create 4 new unique variables that together explain almost 70% of the variance. The number of variables decreased from 15 to 4 which significantly simplified the model. Now, the model could be estimated using these four new variables. However, this thesis will not continue any further as the KFSI are the subject of this thesis and therefore will be kept as variables. Three of four new variables represent a group of KFSI. Moreover, the PC7 was very highly correlated with KFSI 1, 2 and 3 – therefore creating new variable representing the first class of KFSI (Knowledge of beneficial ownership). The PCA provided interesting insight and its results could be used in future for a different analysis.

In summary, the PCA hinted that instead of using 15 components, the model can be simplified by omitting the components and using only the component groups presented in Section 2.3: Knowledge of beneficial ownership; Corporate transparency; Efficiency of tax and financial regulation; and International standards and regulation. The weighting could be done on the basis of group that would be assigned weights instead of components.

5 Methodology

5.1 Survey and weight setting

In order to determine new weights for each of 15 KFSI components, I obtained expert opinions via electronic survey distributed using email. The selection of survey method together with distributing a well-built questionnaire are critical for obtaining a good response rate and data that can be further analyzed. The following chapter deals with selecting a survey method and then describes the questionnaire that was created.

5.1.1 Survey method

In history there have been various survey methods presented. Mail survey and surveys conducted via interview were the most favored ones during the first half of the 20th century. In depth analysis of the development of survey is provided by Converse (2009) that focuses on the development of surveys in the USA until the 1960. However, new technologies such as telephone and fax emerged during the second part of the 20th century that brought new methods of surveys into attention. The end of the century presented new opportunities with development of email and Internet in the world.

The Internet enables collecting large amounts of data with less effort and without even meeting or direct communication with the respondent. It enables to reach respondents from all over the world without the costs of postage, material etc. The analyzing and processing of the answers is also free of errors that can be caused by manual and individual data entry of each answer (Witt and Karlan, 1998). The comparison of the responses obtained by various methods were studied e.g. by Kaplowitz et al (2004) who compared mail and email surveys and found increased response rate for email surveys and recommended email surveys for populations with the access to the Internet. The mail, fax and email survey results were examined by Cobanoglu et al. (2004). Their analysis indicates that web-based survey have significant advantages in terms of response rate and costs. They came to the same results as Keisler and Sproull

(1986) and Bachman, Elfrink and Vazzana (2000). In overall, the web-based surveys are recommended for three reasons. First, they yield higher response rates at lower costs. Second, the majority of the researchers in the USA have the access to the Internet. In present, this can be generalized to the whole world except for a few developing countries. And finally third, that eliminates errors that are caused by rewriting and entering the survey results manually (Cobanoglu et al., 2004).

Greenlaw and Brown-Welty (2009) compared the web-based and paper-based surveys. One of their conclusions highlights the significantly lower costs of web-based surveys over the ones distributed and collected in paper form.

In present, with world's interconnectedness and the Internet connection being a standard in developed countries, the choice of the survey method was simple. The web-based survey distributed via email was selected also due to the above mentioned research and cost benefits it provides.

5.1.2 The structure

When creating a questionnaire and survey that yields good response rate, is clear and at the same time contains all the necessary elements, one must take into account many aspects and decide which form suits his objective the best. One must determine the structure of the survey, evaluate open or closed question approach and for example decide if "I don't know" option will be provided.

In the process of designing the survey I focused primarily on the simplicity and easy comprehensibility to avoid confusion of the respondents that could lead to distortion of the survey results. The simplest approach is also favored by Dillman et al (1998). Furthermore, simple syntax, specific wording and familiar words are preferred (Krosnick and Presser, 2010).

5.1.2.1 Open vs closed

Both open and closed questions have its advantages and disadvantages. Open questions are more costly, have lower response rate and analysis takes longer. On the other hand, in terms of measurement of quantity (in this case – specific weights) they provide more

precise results (Krosnick and Presser, 2010). Closed questions are less costly and easier to analyze and do not demotivate respondents that much. To combine the advantages of both I decided to allow respondent to choose between both of them. This solution was implemented with greater response rate in mind. However, there were no respondents that assess specific weights using open question method and only 4 who provided relevant comments. See more in the section analyzing the results.

5.1.2.2 Scaling points

The number of rating points on the scale differs dramatically across the literature. Carmines, Robinson and Shaver (2001) list range of social psychological surveys and write down the number of scaling points. The results suggest that most preferred number of scaling points is 5, followed by 7 and 4 scaling points, respectively. When choosing the amount of scaling points, it is crucial to focus on their future ease of translation into clear results, clarity of the meaning and the uniformity of the meaning so that every respondent understands it in the the same and there no confusion possible across the respondents (Krosnick and Presser, 2010).

The empirical evidence of using various number of scaling points is examined in the vast literature review provided by Krosnick and Presser (2010). For example, the cross-sectional reliability of 5, 7, 9 and 14 scaling points was proven to be equal by Lissitz and Green (1975). Authors also showed decreasing reliability for declining number of scaling points (3 and 2 scaling points). Similar results were obtained by e.g. Martin (1978) and Srinivasan and Basu (1989). On the other hand the studies on validity of various scaling points show that gains in validity get smaller with increasing number of scaling. This outcome is applicable for minimum two scaling points (Lehmann and Hulbert, 1972; Lissitz and Green, 1975; Ramsay, 1973). Krosnick and Presser (2010) conclude that 7 scaling points are the most convenient but also provide references that 5 scaling points provide similar results in terms of validity, reliability and satisficing. Supported by the emphasis on the simplicity discussed above, I decided to incorporate 5 scaling points to the survey.

These 5 scaling by which the respondent assesses his preferences to each component of the qualitative part were labelled as follows:

- Significantly lower
- Lower
- The same
- Higher
- Significantly higher

Put differently the respondent answers whether he thinks the weight for each KFSI should remain the same of should be (significantly) higher/lower than the current equal weight. As mentioned above, option of open answer was allowed and thus sixth option labelled “Other” was added where the respondent could either assess the specific weights (even 0 if the respondent believed that the component should be omitted from the index) or the respondent could provide comment.

5.1.2.3 Dealing with no opinion answer

Lastly, the dilemma of allowing “I don’t know” or no opinion answer was addressed. There are researches showing that respondent who answered he did not have an opinion in fact did have an opinion that could be relevant answer for the question (Krosnick et al., 2002). Similar results were reached by Gilljam and Granberg (1993) and also by Bradburn and Sudman (1988), Feick (1989) and others. They all came with conclusions that allowing respondent to answer that he does not have an opinion discourages him to think about the answer rather than encourages him to complete the whole survey with better results for the researcher.

On the other hand, many researchers recommend adding this option primarily to avoid non-attitude answers (Krosnick and Presser, 2010). These answers do not truly represent the respondent’s opinion because respondent was forced to answer them even though he might not have any opinion on the matter. Answers like this create noise in the results and it is almost impossible to identify and eliminate them from the results. The proposed solution is allowing respondents to have no opinion. The advocating of this method can be found in Bogart (1972), Vaillancourt (1973) Oppenheim (2000) and

many others. This allows respondents that do not have necessary information or ability to formulate an opinion truly state their inability to answer.

Also, the respondents can have legitimate reasons for not answering the questions or are for some reason unable to answer. Forcing respondents to answer each question mandatorily can in extreme cases cause frustration that can lead to premature termination. In addition, human subject protection committees demand that respondents are told that their answers are voluntary and not forced by any authority (Dillman et al., 1998). For above mentioned reasons I decided to allow respondents to answer “I don’t know” question in spite of the research showing that it does not necessarily provide better results. Moreover, this survey was distributed as a part of larger FSI survey where “I don’t know” answer was provided for questions. To maintain consistency it was only natural to add this option to the survey. In the end, this option was executed only by minority of respondents. See section results for details.

In addition, the 11 principles for constructing web-surveys presented by Dillman et al. (1998) were considered during drafting of the survey. These principles provide general guidelines about coding, format etc.

In summary, the survey is introduced by a short paragraph that briefly explains the current situation of the weighting system, the use of qualitative part and finally gives instructions on how to fill out the survey. The respondent is asked to assess weights to each of the 15 KFSI thus the survey consists of 15 questions only. Each question opens with a short description of KFSI and contains link to further details. In total, the respondent has 7 options for each answer. Five options are to assess his opinion in qualitative manner (significantly lower, lower, the same, higher, significantly higher). One option is to reveal his opinion in quantitative way labelled “Other” where respondent was asked to either insert specific number or zero. Moreover, respondent also could express his opinion in words. The last option was “I don’t know” answer. Its use was justified earlier in the chapter.

All questions were on one webpage and user did not have to confirm every answer. Only one confirmation of the whole survey was required. For the whole survey, refer to Appendix 3.

5.1.3. Survey platform and distribution

The survey and the logic behind it described in the previous sections. The survey was distributed as a part of more extensive FSI Survey that contained other survey questions. The survey used in this thesis represented only one page of the FSI survey.

The FSI survey was designed in English and created using SoSci Survey – an online software allowing to create elaborate survey on their webpage. The survey was sent to 3611 individual email addresses. The email addresses belonged to academics, experts, tax authorities' officials, representatives of non-governmental organizations, and TJN's donors. Along with these, the email was circulated via a few group emailing lists including Financial Transparency Coalition, TJN Europe, TJN Australia and others.

The FSI survey in total comprises of 46 questions but there were filters present so respondent was not necessarily able to access all questions. To prevent multiple entries by single user, each respondent could fill in the survey only once. This does not mean that when respondent did not finish the survey he could not come back to it – it only prevented those who already finished the survey to fill it out again.

The FSI survey introduced the respondent by a welcome screen where instructions and logic of the survey was explained. The FSI survey was targeted on the respondents familiar with the FSI to provide their insights and build on the foundations set by the FSI. Thus the first question asked: "Have you ever heard of the Financial Secrecy Index (FSI)?" If respondent answered "No", the filter was triggered and respondent was presented with following message: "This survey focuses on details of the FSI. Therefore, familiarity with the FSI is essential to answer it. While we kindly appreciate your willingness to fill in the survey, we would suggest you familiarize yourself with the FSI first and then come back to participate in the survey". If the respondent had heard about the FSI, he was allowed to continue with FSI survey. The FSI survey began with few demographic questions and then moved to specific questions about FSI. The

part of the survey relevant for this thesis was presented approximately in the middle of the FSI survey.

The period in which the survey could be filled in began on 18 January 2016 and the deadline was set for 15 March 2016 after which the survey could not be accessed.

5.2 Information theory approach

Information theory and its model evaluation criteria present a different method of quantifying the performance of models. Statistical model selection criteria from information theory combine a component of goodness-of-fit represented by maximum likelihood functions with a component of complexity, which counteracts the first mentioned component. There are numerous criteria rooted in information that each provide different approaches. Thus, the model that is favored by one or more of these statistics can be characterized as parsimonious – archiving a better balance between the goodness-of-fit and the complexity of the model for which it is penalized relative to other models (Stapleton and Garrod, 2007).

In this thesis I create two models (the rationale is explained in the next subsection) and quantify their performance using traditional goodness-of-fit statistics and also model selection criteria from information theory. Concretely, I compare the models using R^2 statistics and Residual Sum of Squares (RSS). The root mean squared deviation (RMSD), Akaike Information Criterion (AIC; Akaike, 1973) and Schwarz Information Criterion (SIC; Schwarz, 1978) take into account not only the goodness-of-fit but also the number of adjustable parameters in models. The Information-Theoretic Measure of Complexity (ICOMP; Bozdogan, 1990) focuses on the complexity of the functional form of the model in addition to taking into account the number of adjustable parameters as a baseline for penalizing goodness-of-fit (Stapleton and Garrod, 2007). Each of the model selection criteria from information theory explains the complexity in the model differently. Thus using a spectrum of these statistics provides a robustness check and also if a model is preferred by multiple information theory model selection criteria, it ensures that the preference of the model was not biased by the disadvantages and preference tendencies that each statistic suffers from. There are more model

selection criteria that use the principles of information theory (e.g. Hannan-Quinn information theory criterion which is similar to AIC and SIC, and also Minimum Description Length (MDL), which on the other hand takes into account the complexity of the function form similarly to ICOMP, and many others). These criteria were not performed due to the fact that they are statistically and computationally demanding and a criterion similar to them is already used.

The following equations represent each information criterion used in the thesis and the show how each statistics was computed.

$$RMSD = \sqrt{\frac{SSE}{n-k}} \tag{5}$$

$$AIC = -2\ln(ML) + 2k \tag{6}$$

$$SIC = -2\ln(ML) + kln(n) \tag{7}$$

$$ICOMP = -2\ln(ML) + kln\left(\frac{tr(\Omega(\theta))}{k}\right) - ln|\Omega(\theta)| \tag{8}$$

Where *SSE* is the model sum of squares error; *n*, the number of data points; *k*, the number of free adjustable parameters; ML, the maximized likelihood function; *Ω(θ)*, the covariance matrix of parameter estimates and *tr(…)* represents trace of matrix.

From the equations above it is clear that the negative part of the equation is the one providing the goodness-of-fit and the positive part penalizes the complexity based on the number of free adjustable parameters and the function form. Consequently, the model which is awarded the lower score is favored. This notion holds for each of the information theory criteria. It is essential to note that this conclusion holds universally. In other words, even if the scores of two comparing models are both negative, based on the equations, the lower score is still preferred and not the score closer to zero.

Chowdhury and Squire (2006) conducted an opinion survey to set a new weighting system for HDI and CDI, which both function under the assumption of equal weight for their components. The results of the survey suggested that there was little justification for relaxing the equal weights for all three components of HDI, as the new

weights derived from the survey were not significantly different from the equal weights. On the other hand, the results for CDI showed that new weights for four out of a total of six components were significantly different from the equal weights. Later, the new weighting system was examined by the information theory approach of Stapleton and Garrod (2007) for HDI and Stapleton and Garrod (2008) for CDI. This thesis follows the main structure and methodology that is shared with both mentioned papers. However, there are few distinctions between the two that will be covered in the relevant section. This thesis finds a balance between the two papers and carefully evaluates which approach better suits the needs arising from the results of survey on qualitative components of FSI.

The following equation (9) shows the sum of secrecy credits for all KFSI components awarded to each particular jurisdiction and serves as a baseline for the computation of the overall secrecy score (the secrecy score is the percentage of secrecy credits obtained out of the maximum amount of 15 – recall equation 3). This shows the current state of the equal weights of each component. The equation X in fact also represents a model that fits the data perfectly. Put differently, if this model is tested and an OLS regression is run, it would show that each coefficient is significant and equal to 1 and the R^2 of the model would be equal to 1 as well.

$$Secrecy\ credits_i = \sum_{j=1}^{15} KFSI_{ij} \tag{9}$$

Where *j* represents each of fifteen KFSI for *i*-th jurisdiction.

There is no method of measuring these components in real conditions that are in fact a theoretical constructs or any scale with units that would allow measuring and comparing the values of different jurisdictions. Thus the dataset that would contain these real observations and could be used to test how well the theoretical model (equation 9) explains the reality simply cannot be constructed. The assumption that “real” secrecy scores lie about unequally weighted components is made (Stapleton and Garrod, 2007; Stapleton and Garrod, 2008). This way the equation 9 does not represent a model that fit the data perfectly.

5.3 Formulating models and generating datasets

Two models were formulated. Model 1 represents the current state of equal weights where all coefficients are constrained to be equal. Thus, the first model has only one parameter (labeled as *k* in the equations 5-8), which does not lead to great penalization in the model selection criteria. Model 2 allows each coefficient to be adjustable under the model fitting. Necessarily, this Model 2 will provide better goodness-of-fit than the constrained model under the equal weighting assumption. On the other hand, the number of adjustable parameters is fifteen instead of one. The question is whether the higher goodness-of-fit will be counteracted by the complexity penalization and ultimately which model of the two will be considered as parsimonious.

Model 1:

$$SC = \beta kfsi1 + \beta kfsi2 + \beta kfsi3 + \beta kfsi4 + \beta kfsi5 + \beta kfsi6 + \beta kfsi7 + \beta kfsi8 + \beta kfsi9 + \beta kfsi10 + \beta kfsi11 + \beta kfsi12 + \beta kfsi13 + \beta kfsi14 + \beta kfsi15 \quad (10)$$

Where *SC* is an abbreviation of Secrecy credits and β coefficient is the same for all fifteen variables.

Model 2:

$$SC = \beta_1 kfsi1 + \beta_2 kfsi2 + \beta_3 kfsi3 + \beta_4 kfsi4 + \beta_5 kfsi5 + \beta_6 kfsi6 + \beta_7 kfsi7 + \beta_8 kfsi8 + \beta_9 kfsi9 + \beta_{10} kfsi10 + \beta_{11} kfsi11 + \beta_{12} kfsi12 + \beta_{13} kfsi13 + \beta_{14} kfsi14 + \beta_{15} kfsi15 \quad (11)$$

Where *SC* is an abbreviation of Secrecy credits and β_1, …., β_{15} are adjustable coefficients under the model fitting.

Stapleton and Garrod (2008), when analyzing the new weights for CDI, constructed a third model where the coefficients for components for which the new weights were significantly different from equal weights, were adjustable under model fitting. The coefficients for components for which the new weights were not significantly different from equal weights were constrained to hold the equal weight assumption. The

assumption of the sum of all weights being equal to one is also enforced. In this thesis creating such a model would not make any sense. There is only one component for which the new weights were significantly different from the equal ones. Under the constraint of the sum of all the coefficients being equal to 1 (in the case of this thesis it would be 15) the model would be the same as Model 2 since the only coefficient being able to be adjustable would in fact not be adjustable due to the constraint of other variables.

Referring to the previous assumption made about the "real" secrecy score, the three datasets were generated in order to test the two constructed models. Similarly to Stapleton and Garrod (2008), the "real" secrecy scores depicted in the three datasets, were each allowed to vary differently around the proposed weights from the conducted survey. The construction of the dataset used a formula that is similar to Model 2 but the coefficients were randomly generated around the new weights.

$$Dataset = \alpha_1 kfsi1 + \alpha_2 kfsi2 + \alpha_3 kfsi3 + \alpha_4 kfsi4 + \alpha_5 kfsi5 + \alpha_6 kfsi6 + \alpha_7 kfsi7 + \alpha_8 kfsi8 + \alpha_9 kfsi9 + \alpha_{10} kfsi10 + \alpha_{11} kfsi11 + \alpha_{12} kfsi12 + \alpha_{13} kfsi13 + \alpha_{14} kfsi14 + \alpha_{15} kfsi15 \quad (12)$$

Where *kfsi1, …, kfsi15* are the data for all jurisdictions obtained from the FSI and discussed in Section 4. And $\alpha_1, \ldots, \alpha_{15}$ are generated as follows:

The "real" data that serve as dependent variables in the models were allowed to randomly vary around the means (new proposed weights) by 10%, 20% and 30%, respectively, and multiplied by the score from FSI. Thus the dataset was allowed to vary 10% around the weights for the least variable and the dataset allowed to vary 30% around the weights is the most variable.

The following Table 5 depicts how three mentioned datasets were generated graphically.

Table 5: The generating of the three datasets

$\alpha_1 =$	1,054	+/-	10%/20%/30%
$\alpha_2 =$	1,016	+/-	10%/20%/30%
$\alpha_3 =$	1,042	+/-	10%/20%/30%
$\alpha_4 =$	0,995	+/-	10%/20%/30%
$\alpha_5 =$	0,954	+/-	10%/20%/30%
$\alpha_6 =$	1,038	+/-	10%/20%/30%
$\alpha_7 =$	1,030	+/-	10%/20%/30%
$\alpha_8 =$	0,959	+/-	10%/20%/30%
$\alpha_9 =$	0,944	+/-	10%/20%/30%
$\alpha_{10} =$	1,009	+/-	10%/20%/30%
$\alpha_{11} =$	1,022	+/-	10%/20%/30%
$\alpha_{12} =$	1,036	+/-	10%/20%/30%
$\alpha_{13} =$	0,896	+/-	10%/20%/30%
$\alpha_{14} =$	0,974	+/-	10%/20%/30%
$\alpha_{15} =$	1,030	+/-	10%/20%/30%

To summarize, two complex models will be tested against three datasets. The three datasets that serve as a dependent variable and under the assumption depict the “real” observed secrecy credits to test the model were constructed around the newly proposed weights and multiplied by the credits they were awarded by FSI. The independent variables were the FSI data for all jurisdictions discussed in the Section 4. The next Section 6 presents the results obtained following the presented methodology.

6 Results

6.1 Results of the survey

Out of the large number of emails sent to solicit responses for the FSI survey, only 86 filled out the entire FSI survey. Another 124 respondents filled out the survey partially. This information concerns the whole FI survey. However, since the following paragraph, only the part relevant for the thesis is discussed.

The total number of 99 responses was received from 43 countries. First, these raw data had to be filtered. Eighteen responses had to be omitted because the respondent did not get that far in the survey or answered “I don’t know” for all of the 15 components, indicating that that the respondent had no intention of filling out the survey. One respondent used the open answers for all of the 15 questions, in which he recommended scrapping the arbitrary weighting of various index components and instead to try to aggregate indicators with PCA. This recommendation is addressed in this thesis and PCA is performed (see the relevant chapter for details). In addition, four respondents used the open answer space to assign zero weight to a total of seven of the components (three respondents suggested each that one of the components should be omitted, and one respondent suggested three components should be eliminated from the index). The remaining majority of the respondents answered via closed answers.

The new weights for 15 components are assessed from the remaining 80 responses. These responses were provided by addressees from 39 jurisdictions. The most responses hailed from the United Kingdom (12), followed by Germany (7) and Switzerland (5). For the table depicting nationalities of all respondents refer to Appendix 4. The most responses arrived from the developed OECD countries, usually from Europe. This composition of population is caused by the fact that most of the emails with survey attached were sent out to addresses from these particular locations. The average respondent was 49,8 years old, and on average had heard about the FSI

for the first time around 3,7 years ago. In terms of the respondents' background, most of the respondents (almost 30%) were from non-governmental organizations. About 23% of respondents were academics or professors, 17% consider themselves experts and 17% work for public services. The remaining respondents were students, journalists or other. In total, 1083 individual questions detecting the opinion on weight of particular components were answered. Another 117 were answered "I don't know" which is less than 10% of the total responses. That means that on average, each of 80 respondents assessed weights to 13,54 components out of 15.

The respondents were the most decisive about KFSI 1 (Banking Secrecy), KFSI 6 (Country-by-country reporting) and KFSI 8 (Efficiency of tax administration) with 75 individual responses and were least conclusive about KFSI 4 (Public company ownership) and KFSI 9 (Avoids promoting tax evasion) with just 69 answers each.

The "I don't know" answers were evaluated as blank answers thus they were not included into the calculation of new weights. However, I believe that including this option to the survey reduced the noise in the results.

The following Table 6 shows each component with the corresponding number of responses it received from each scaling point. It is clear from the table that higher weights were assessed in 432 cases opposed to lower or zero weights that were chosen by respondents only 104 times. Keeping the equal weights was in total selected by 875 individual responses. In aggregate, higher weights were preferred more than 4x more frequently than lower or zero weights. This hints that all components should be weighted higher than they were before. This is tackled at the beginning of the next chapter.

Similar phenomenon can be observed in other surveys. Krosnick (1991) noted that respondents usually tend to aim at satisfying the researcher instead of optimizing their responses. Evidence on surveys which adopt agree/disagree or yes/no answers shows that positive answer (agree or yes) is chosen more frequently than a negative one (Berg and Rapaport, 1954). Krosnick and Presser (2010) label this issue as "acquiescence." They present further empirical research supporting these results and providing

additional evidence. This could potentially be the cause of the preference of higher weights over lower weights as an alternative to yes/no answer. However, the proportion of the ratio is not consistent with the evidence introduced in Krosnick and Presser (2010).

Nevertheless, this fact does not have any effect on the weights as they are normalized to provide easily comprehensible results. The next chapter provides more information on the matter.

Table 6: Number of different responses for each component

	I don't know	**Zero**	**Sign. lower**	**Lower**	**The same**	**Higher**	**Sign. higher**
KFSI 1	5	0	2	1	36	25	11
KFSI 2	8	0	2	4	39	15	12
KFSI 3	7	0	1	5	30	29	8
KFSI 4	11	0	1	4	40	17	7
KFSI 5	8	0	1	8	43	14	6
KFSI 6	5	1	2	4	31	25	12
KFSI 7	7	0	0	4	39	21	9
KFSI 8	5	2	1	6	40	20	6
KFSI 9	11	1	4	4	39	14	7
KFSI 10	9	0	2	3	38	20	8
KFSI 11	9	0	0	5	353	25	6
KFSI 12	8	0	0	5	38	27	12
KFSI 13	9	0	4	12	37	14	4
KFSI 14	8	0	2	9	34	20	7
KFSI 15	7	0	1	3	38	22	9
total	117	4	23	77	875	308	124

Source: the survey results

Four main potential sources of error that arise from surveying a population have to be taken into account – Coverage error, sampling error, measurement error and nonresponse error (Groves, 1991).

The survey was aimed only at individuals familiar with the FSI. The respondents were asked if they know about FSI. There was no test to verify if this statement was true. Therefore, there is a possible problem of self-selection bias. There also can be sampling bias, response and non-response biases present.

6.2 KFSIs under the new weighting system

The previous paragraphs described the process of filtering the responses and the reasoning for doing so. First, it was important to quantify the articulated options into real weights. The inclination towards simplicity is repeated during the survey and it is an important part of the decision process. Thus, I followed the simplest and the most intuitive arrangement to quantify the weights. Naturally, if the respondent awarded zero weight to components, it received 0 weight. For each of 5 scaling points, equivalent weigh was awarded. For “Significantly lower” answer, a weight equal to 1 was awarded; weight equal to 2 was awarded for “Lower” answer; and a response of “The same” received a weight equal to 3. Similarly, the weight 4 was assigned to “Higher” result, and finally, the “Significantly higher” response obtained a weight of 5.

This method was chosen due to its simplicity and intuitiveness, as was mentioned earlier. Different evaluation methods were considered and performed as well. The greatest drawbacks were that these methods make additional assumptions, and more importantly, the weights proposed by these methods were not significantly different from the weights proposed by the intuitive approach. See Section 6.3 for details of the testing.

As already mentioned in the previous section, the means of weights for all components were larger than equal weights result for each component. To set new weights I computed the means for all 15 components. In order to follow the FSI methodology, the new weights were adjusted so that their sum would be equal to 15. This arrangement rescales the weights in a way such that the component obtaining the lowest weight from the survey would have a weight that is lower than the equal weight assumption. This helps to address the acquiescence issue discussed at the end of the previous chapter. Recall using a budget for setting weights from earlier in the thesis, this is an application of the method in order to normalize the weights because increasing all the weights at the same time would not make empirical sense.

Table 7: Weights for each component derived from the survey

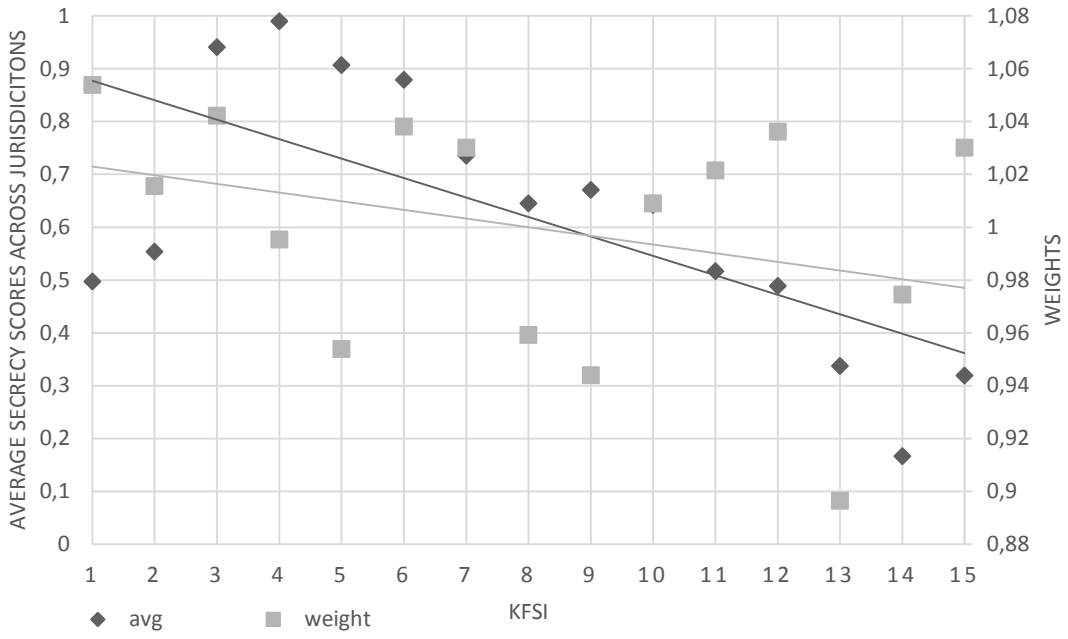
	Weight	**SE**	**SD**	**rank**
KFSI 1	1,054	0,029	0,254	1
KFSI 2	1,016	0,033	0,276	8
KFSI 3	1,042	0,029	0,247	2
KFSI 4	0,995	0,029	0,238	10
KFSI 5	0,954	0,028	0,239	13
KFSI 6	1,038	0,034	0,298	3
KFSI 7	1,030	0,027	0,232	5
KFSI 8	0,959	0,033	0,283	12
KFSI 9	0,944	0,036	0,297	14
KFSI 10	1,009	0,030	0,253	9
KFSI 11	1,022	0,026	0,223	7
KFSI 12	1,036	0,030	0,253	4
KFSI 13	0,896	0,032	0,269	15
KFSI 14	0,974	0,032	0,270	11
KFSI 15	1,030	0,028	0,242	5

Source: own estimation; Stata software output

The Table 7 present the weights for all fifteen components of the qualitative part of the FSI that I determined from the results of conducted survey. The first column shows the derived weights, the second column displays standard errors of respective components. The third column presents standard deviations calculated from the survey results and the last column shows the rank of the individual components. The KFSI 1 which stands for Banking Secrecy was awarded the greatest weight by the respondents of the survey. The recorded ownership which stands behind the KFSI 3 ranked second and the KFSI 6 (Country-by-county reporting) obtained third greatest weight. On the other hand, the KFSI 13 which is labeled Bilateral Treaties obtained the lowest weight based on the opinions of the respondents.

In addition, I plotted these newly proposed weights against the averages of the KFSIs that were awarded for each jurisdiction (similar to the descriptive statistics in Section 4) to determine whether the opinions of the respondents were correlated with the real observations. The pair-wise correlation between the two variables is 0,081, showing that they are not correlated, a finding that is also observable from Figure 2.

Figure 2: Weights for each component derived from the survey plotted with average values of secrecy scores for each KFSIs across jurisdictions



Source: Financial Secrecy Index 2015; own estimation

I tested the hypothesis that the new weights are statistically the same as the equal weights. I used a t-test and compared equal weight (equal to 1) to normalized mean (the new weight) computed from the normalized results of the survey for each component. Surprisingly, the hypothesis was rejected at the 95% confidence interval in only one case – for KFSI 13 (see output from statistical software Stata in Appendix 5). This means that the hypothesis could not be rejected for the remaining 14 components and only the thirteenth component's new weight was found to be significantly different from the equal weight. Formally, I will refer to these results that the weights of 14 components were not significantly different from the equal weights configuration.

Ranking of the jurisdictions based on the equal weights and also for the newly proposed weights was constructed for secrecy scores as well as for the aggregate FSI. As expected, there were only minor changes in the rankings. For both secrecy scores and FSI the ranking of the top 20 jurisdictions did not change. To illustrate the change,

Switzerland occupied the 32nd position of secrecy scores under the equal weights configuration and moved to the 33rd position under new weighting system. The Seychelles on the other hand moved up from 33rd to 32nd position. There are several other similar cases but there were no dramatic changes in the ranking of the jurisdictions. The Spearman rank order correlation coefficient was computed for rank derived from equal weight and rank constructed from proposed weights. It amounted to 0,9997 for both the ranks of secrecy scores and FSI. In addition, the correlation coefficient using Kendall's rank correlation coefficient for secrecy scores rank was 0,9918 and 0,9922 for the FSI ranking. The pair-wise correlation between secrecy scores created under equal weights and those computed under the new weighting system was 0,9999. The same results were recorded for the FSI score under these two mentioned weights.

Surprisingly, the new weights that I derived from the survey were not significantly different from the currently used equal weight for 14 out of 15 components. Importantly, there was no evidence from the survey that any component should be omitted from the index and all of them were found to be valid. In addition, the new weights did not provide scores that were significantly different from those resulting from the equal weights configuration. Finally, there were no dramatic changes in the ranking of jurisdictions for secrecy scores or the composite FSI index. These results provide support for keeping the equal weighting system rather than implementing one that was derived from the conducted survey. The Chapter 6.4 applies the information theory approach on the proposed weights in this Chapter 6.1 and decides if using the simple model of equal weights is justified and preferred to using the more complex model of unequal weights.

6.3 Another evaluation methods

The results obtained from the survey had to be evaluated in order to transform them into weights. Earlier, I explained the intuitive method that I used and presented the results. However, I also evaluated the answers differently to see how the weights would change under different evaluation of respondents' answers.

The following Table 8 shows how I alternatively evaluated the responses from the survey. The intuitive weights that I used in the thesis are in the first column. Unlike the other weights that I tested, the intuitive evaluation is not anchored around 1. The opinion that the weights should remain the same is evaluated as 1 in the other tested in other evaluation methods but not in the intuitive one – it is evaluated as 3. But since all the weights are normalized so that their sum is fifteen, this does not have any effect.

Columns labeled Weights 1 to Weights 4 represent the alternative evaluation systems. Each new evaluation was created by multiplying and dividing the equal weight option by two numbers. In other words, for example for Weights 1, I multiplied "The same" opinion labeled by 3 in the survey by 1,5 and 2, respectively, to evaluate "Higher" and "Significantly higher" opinions. Similarly, I divided "The same" opinion by 1,5 and 2 to evaluate "Lower" and "Significantly lower" responses. The identical approach was applied to Weights 2 to Weight 4. The number that were used to multiply/divide is clear from the Table 8. In the most extreme case of Weights 4, I assume that if the respondent answered "Significantly higher/lower," he in fact believed that the weights should be ten times higher/lower than equal weights and five times higher/lower if he answered "Higher/Lower" which is a very strong assumption.

Table 8: Alternative methods of evaluating the answers from the survey

Answer in survey	Intuitive	Weights 1	Weights 2	Weights 3	Weights 4
0	0	0	0	0	0
1	1	0,5	0,25	0,16	0,1
2	2	0,66	0,5	0,33	0,2
3	3	1	1	1	1
4	4	1,5	2	3	5
5	5	2	4	6	10

Source: own computation

The new evaluation systems assume more variable opinions than the intuitive one. The less variable evaluation methods were not considered since they would necessarily provide the same results in terms of significance. In other words, 14 or 15 weights would not be significantly different from the equal weights. The aim of assuming different evaluation methods was to determine if the weights would change in terms of

their significant difference from the equal weights – therefore assumptions about more variable evaluations were made.

The weights were derived from the evaluations by the same technique as the intuitive weights. Following Table 9 represents weights assessed based on the previously made evaluation assumptions presented in Table 8.

Table 9: Weights derived from alternative evaluation methods

	Intuitive	**Weights 1**	**Weights 2**	**Weights 3**	**Weights 4**
kfsi1	1,054	1,068	1,118	1,155	1,188
kfsi2	1,016	1,021	1,063	1,068	1,072
kfsi3	1,042	1,054	1,077	1,114	1,152
kfsi4	0,995	0,987	0,967	0,950	0,932
kfsi5	0,954	0,937	0,883	0,839	0,797
kfsi6	1,038	1,057	1,123	1,168	1,213
kfsi7	1,030	1,029	1,044	1,051	1,056
kfsi8	0,959	0,951	0,916	0,898	0,881
kfsi9	0,944	0,940	0,909	0,883	0,857
kfsi10	1,009	1,010	1,010	1,012	1,012
kfsi11	1,022	1,020	1,006	1,014	1,022
kfsi12	1,036	1,040	1,089	1,102	1,113
kfsi13	0,896*	0,882*	0,794*	0,739*	0,693*
kfsi14	0,974	0,973	0,952	0,944	0,939
kfsi15	1,030	1,033	1,050	1,063	1,074

* Significance of difference from equal weights at the 95 % confidence level
Source: own estimation; Stata software output

First, I tested whether the new weights are equal to the weights proposed by the thesis by intuitive approach. For Weights 1, this hypothesis could not be rejected for all 15 components at 95% confidence interval which suggests that the weights are almost identical to the intuitive ones. The results for more variable Weights 2-4 show that only KFSI13 is significantly different from the suggested intuitive weights.

In addition, I hypothesize that the newly suggested weights are equal to ones equal to one (equal weights assumption). Surprisingly, the results of t-tests showed the same results across Weights 1 – Weights 4. The hypothesis was rejected only for KFSI13 in all four weighting schemes. The hypothesis could not be rejected for the remainder of 14 variables at 95% confidence.

The results proved that using the simpler, intuitive evaluation is justified. Making other assumptions about the opinions of the respondents did not produce significantly different weights. The newly suggested weights were not significantly different from the equal weights configuration in 14 out of 15 components. This shows that the weights are very similar regardless of the evaluation method (even when creating strong assumptions about the respondents' opinions).

6.4 Results of Information theory approach

In the Chapter 6.2, I focused on constructing a new weighting system for the 15 components of qualitative part of the FSI. I conducted an opinion survey and I derived new weights for each component based on the survey results. In this chapter, I determine whether using a more complex model featuring unequal weights is justified compared to using the much simpler model set up under the assumption of equal weights.

The Table 10 shows the results of testing each model against 3 datasets. Each dataset was evaluated separately. Statistics that strike better goodness-of-fit or model selection criteria are highlighted in bold in respect with each model and each statistics. Model that minimize RSS, RMSD, AIC, SIC and ICOMP and maximize the R^2 is preferred.

Table 10: Goodness-of-fit and models selection statistics from information theory for two formulated models tested on generated hypothetical datasets and updated FSI 2015

Model	Dataset	RSS	R^2	RMSD	AIC	SIC	ICOMP
1	1	3,176	0,9997	0,1773	-62,414	**-59,789**	-32,207
2	1	**2,195**	**0,9998**	**0,1588**	**-72,093**	-32,718	**-33,350**
1	2	12,0202	0,9987	**0,3449**	**73,348**	**75,973**	**35,674**
2	2	**11,1247**	**0,9988**	0,3579	93,450	132,820	49,410
1	3	27,8324	0,9971	**0,5249**	**158,988**	**161,614**	**78,494**
2	3	**24,5823**	**0,9974**	0,5315	174,323	213,698	89,849

Source: own estimation; Stata software output

Model 2 provides a better fit for Dataset 1 than Model 1. The goodness-of-fit statistics RSS and R^2 as well as RMSD, AIC and ICOMP from the information theory model

selection statistics favor the more complex model. The number of free parameters is penalized more by the SIC compared to AIC which explains why AIC and SIC both favor different models. The difference between the models' ICOMP model selection statistics is small, which hints that the models are performing similarly and the more complex model is preferred only slightly over the simple one. However, ICOMP as well as other model selection statistics is unitless and thus the numbers cannot be evaluated separately, so the conclusion can only be derived from the comparisons of different models. Put differently, if a model scored 1000 from ICOMP statistics, no conclusion can be made until it is put into perspective with the score of another model. The only exception is R^2: you can evaluate a model by its value and do not necessarily have to compare it with other models.

For Dataset 2, all model selection statistics from information theory favor the simpler model. Unsurprisingly, the goodness-of-fit statistics show that Model 2 provide a closer fit for Dataset 2.

Finally, for Dataset 3 the preference is identical to the one of Dataset 3 – all information theory model selection criteria favor Model 1 using equal weights. Similarly, the goodness-of fit statistics of Model 2 with coefficients adjustable under model fitting show that the more complex Model 2 provides a better fit.

The values of model selection statistics are increasing with the variability of the dataset. Also, the R^2 is decreasing, in line with expectations. Naturally, the model will be a better fit for a relatively more consistent dataset rather than for a more variable dataset. As with the model selection statistics, the goodness-of-fit component of the statistics is decreasing with the complexity of the model and the penalization.

Generally, the simpler Model 1 is favored by nine information theory model selection criteria out of a total of twelve. This reveals a general preference for the simpler model due its low complexity and low number of free parameters. Interestingly, the more complex Model 2 was selected only for the least variable Dataset 1. Low variability causes the generated data to tend more towards the means and the Model 2 then fits the data almost perfectly. We also tested a lower variability of 5% and 2% around the

weights and the results show also the preference of more complex model by the selection criteria. This is caused by the simple fact that in these cases the variability of the data is so low that the model with adjustable coefficients explains the data almost perfectly (R^2 converges to 1).

The fact that Model 1 is favored over Model 2 for more variable datasets can be only partially explained by collinearity (Table 11). The majority of the variables are not highly collinear but some collinearity can be observed. If the variables would be highly collinear (in other words the increase in one variable would correspond to an increase in all other variables), then Model 2 would be even less likely to be favored by the model selection statistics. In the extreme case, if the variables would be perfectly collinear, the differentiating of the weights would not make any sense (Stapleton and Garrod, 2008).

Table 11: Pearon's correlation matrix for all KFSI components

KFSI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1,000														
2	0,006	1,000													
3	0,297	0,122	1,000												
4	0,268	-0,102	0,116	1,000											
5	0,232	0,068	0,682	0,074	1,000										
6	0,312	0,126	0,992	0,125	0,676	1,000									
7	0,427	0,283	0,212	0,132	0,254	0,330	1,000								
8	0,534	0,321	0,260	0,181	0,346	0,474	0,763	1,000							
9	0,254	0,253	0,237	-0,023	0,098	0,235	0,454	0,473	1,000						
10	0,096	0,416	0,237	-0,010	0,161	0,258	0,399	0,383	0,444	1,000					
11	0,538	-0,010	0,334	0,211	0,186	0,339	0,156	0,324	0,050	-0,049	1,000				
12	0,463	0,020	0,582	0,166	0,421	0,578	0,302	0,386	0,111	-0,069	0,391	1,000			
13	0,567	0,061	0,448	0,210	0,350	0,485	0,286	0,485	0,166	-0,074	0,450	0,759	1,000		
14	0,432	0,168	0,582	0,171	0,448	0,594	0,419	0,594	0,256	0,254	0,246	0,410	0,425	1,000	
15	0,408	0,031	0,299	0,276	0,213	0,306	0,202	0,308	0,091	-0,044	0,742	0,361	0,436	0,260	1,000

Source: own estimation; Stata software output

The results of the information theory approach applied to the new weights of the qualitative components of FSI derived from the opinion survey presented in this thesis show that there is little justification for relaxing the equal weights assumption. The complexity of the model with unequal weights counteracts the goodness-of-fit. These results are in line with those Stapleton and Garrod (2007) published about newly

proposed weights on HDI (Chowdhury and Squire, 2006) and those by Stapleton and Garrod (2008) presented about the new weighting system of CDI. In both papers they conclude that the increased goodness-of-fit of the more complex models does not justify the increased complexity and the equal weighting system should not be abandoned since it the models using equal weights are parsimonious.

7 Conclusion

Financial secrecy is a dangerous developing issue that affects not only developing countries but also the developed economies of the world, as was well depicted in Chapter 1. Identifying the conduits and recipients of illicit financial flows is crucial for addressing this issue effectively. The Financial Secrecy Index ranks jurisdictions by their secrecy (qualitative part) and their offshore financial activities (quantitative part). The qualitative part of the index is composed of 15 components divided into 4 categories that are all equally weighted, a design that is generally subject to criticism. This thesis aims at setting new weights for each component and tests the statistical validity of the new, unequal weighting system.

By conducting an opinion survey I assessed a new weighting system based on the results of the survey and tested if the newly proposed weights differ statistically from the equal weights configuration. The hypothesis that the new weights are the same as the equal ones could not be rejected in 14 out of 15 components based on the conducted opinion survey and assessed weights.

In addition, the thesis builds on the published application of information theory criterions (Stapleton and Garrod, 2007; Stapleton and Garrod, 2008) on the new weights for HDI and CDI (Chowdhury and Squire, 2006) and applies it to the 15 qualitative components of FSI to address criticisms related to the use of equal weights. In order to test information theory criteria I generated multiple datasets that were allowed to vary differently around the newly proposed weights and tested them against the equally weighted and unequally weighted FSI set-ups. Both models are run using data on 102 jurisdictions from the results of the qualitative part of the latest FSI 2015.

As expected, the more complex model provides superior goodness-of-fit statistics over the model using equal weights. However, the simpler model with equal weights is in the majority of cases favored by the model selection statistics rooted in information

theory. Put differently, the better goodness-of-fit is in overall counteracted by the penalization of the complexity of the model originating in information theory and in general does not provide justification for the relaxation of the equal weighting system.

This thesis concludes with a descriptive and statistical analysis of the dataset that was performed to identify any statistical relationships that would support using an alternative weighting system, but none were found. However, the results of the principal components analysis hints that the 4 classes into which the qualitative components are grouped could serve a baseline for weighting instead of the 15 individual components.

In summary, this thesis aimed at proposing new weights for the 15 components of the qualitative part of the FSI to replace the current equal weighting system. However, the newly proposed weights from the opinion survey were not found to be significantly different from the equal ones in 14 out of the 15 components, and further analysis using model selection criteria from information theory favored the simple model of equal weights. Additionally, the rank of jurisdictions in the FSI was not significantly different from the one using equal weights.

Based on these results, the thesis suggests that the equal weight approach is preferred statistically over the unequal weight approach, in spite of criticisms, and that the weights for the qualitative components of the FSI should not diverge from the equal weights assumption. For future research, more emphasis could be put on simplifying the qualitative part of the FSI. New specification of the components, maybe using only 4 groups as was suggested by PCA analysis, could yield interesting results as the number of parameters would decrease significantly and an opinion survey could result in proposing weights significantly different from the equal weights. Additionally, applying an information theory approach would penalize the complexity of the model to a lesser extent, and a different model could be favored.

Bibliography

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Proceedings of the Second International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 267–281.

Anand, S. and Sen, A., 1997. Concepts of human development and poverty: a multidimensional perspective..Human development papers, Human Development Report Office (UNDP), New York.

Arrow, K.J., 1950. A difficulty in the concept of social welfare. The Journal of Political Economy, pp.328-346.

Bachmann, D.P., Elfrink, J. and Vazzana, G., 1999. E-mail and snail mail face off in rematch. Marketing Research, 11(4), p.10.

Baker, R., 2007. The ugliest chapter in global economic affairs since slavery.*Global Financial Integrity, New York*.

Becker, G.S., Philipson, T.J. and Soares, R.R., 2003. The quantity and quality of life and the evolution of world inequality (No. w9765). National Bureau of Economic Research.

Berg, I.A. and Rapaport, G.M., 1954. Response bias in an unstructured questionnaire. *The Journal of Psychology*, *38*(2), pp.475-481.

Bogart, L., 1972. Silent politics: Polls and the awareness of public opinion. New York: Wiley-Interscience.

Boyce, J.K. and Ndikumana, L., 2012. Capital Flight from Sub-Saharan African Countries: Updated Estimates, 1970-2010. *PERI Working Papers*.

Bozdogan, H., 2000. Akaike's information criterion and recent developments in information complexity. Journal of Mathematical Psychology 44, 62–91.

Bozdogan, H., 1990. On the information-based measure of covariance complexity and its application to the evaluation of multivariate linear models.*Communications in Statistics-Theory and Methods*, *19*(1), pp.221-278.

Bradburn, N.M. and Sudman, S., 1988. Polls & surveys: understanding what they tell us (Vol. 41). Jossey-Bass Inc Pub.

Brandolini, A., 2008. On applying synthetic indices of multidimensional well-being: Health and income inequalities in selected EU countries. Bank of Italy Temi di Discussione (Working Paper) No, 668.

Carmines, E.G., Robinson, J.P., Shaver, P.R. and Wrightsman, L.S., 2001. Measures of Political Attitudes.

Carson, R.T. and Louviere, J.J., 2011. A common nomenclature for stated preference elicitation approaches. Environmental and Resource Economics,49(4), pp.539-559.

Cobanoglu, C., Warde, B. and Moreo, P.J., 2001. A comparison of mail, fax and web-based survey methods. International journal of market research,43(4), pp.441-452.

Cobham, A., 2005. Tax evasion, tax avoidance and development finance.*Queen Elizabeth House, University of Oxford*, *129*(4).

Chowdhury, S., and Squire, L., 2006. Setting weights for aggregate indices: An application to the commitment to development index and human development index. Journal of Development Studies, 42(5), 761–771. doi:10.1080/00220380600741904

Cobham A., Janský P. and Meinzer M., 2015. The Financial Secrecy Index: Shedding New Light on the Geography of Secrecy. CGD Working Paper 404. Washington, DC: Center for Global Development. http://www.cgdev.org/publication/financial-secrecy-index-shedding-new-light-geography-secrecyworking-paper-404

Converse, J., 2009. Survey Research in the United States: Roots and Emergence, 1890–1960. Piscataway, NJ: Transaction Publishers

De Kruijk, H. and Rutten, M., 2007. Weighting dimensions of poverty based on peoples priorities: Constructing a composite poverty index for the Maldives. *Q-squared working paper*, *35*.

Decancq, K. and Lugo, M.A., 2010. Weights in multidimensional indices of wellbeing: An overview. Econometric Reviews, 32(1), pp.7-34.

Desai, M. and Shah, A., 1988. An econometric approach to the measurement of poverty. *Oxford Economic Papers*, *40*(3), pp.505-522.

Dillman, D.A., 2000. Mail and internet surveys: The tailored design method(Vol. 2). New York: Wiley.

Dillman, D.A., Tortora, R.D. and Bowker, D., 1998, August. Principles for constructing web surveys. In Joint Meetings of the American Statistical Association.

Feick, L.F., 1989. Latent class analysis of survey questions that include don't know responses. Public Opinion Quarterly, 53(4), pp.525-547.

Financial Secrecy Index, 2015. Financial Secrecy Index 2015 Methodology. *Tax Justice Network.*

Fleurbaey, M. and Gaulier, G., 2009. International Comparisons of Living Standards by Equivalent Incomes*. *The Scandinavian Journal of Economics*, *111*(3), pp.597-624.

Fusco, A. and Dickes, P., 2008. The Rasch model and multidimensional poverty measurement. In *Quantitative approaches to multidimensional poverty measurement* (pp. 49-62). Palgrave Macmillan UK.

Gilljam, M. and Granberg, D., 1993. Should we take don't know for an answer?. Public Opinion Quarterly, 57(3), pp.348-357.

Greenlaw, C. and Brown-Welty, S., 2009. A comparison of web-based and paper-based survey methods testing assumptions of survey mode and response cost. Evaluation Review, 33(5), pp.464-480.

Groves, R.M., 1991. Measurement error across disciplines. Measurement errors in surveys, pp.1-25.

Henry, J.S., 2012. The price of offshore revisited. *Tax Justice Network, http://www. taxjustice. net/cms/upload/pdf/Price_of_Offshore_Revisited_120722. pdf.*

Hotelling, H., 1933. Analysis of a Complex of Statistical Variables into Principal Components. Journal of Educational Psychology , 24 (6 & 7), 417–441 & 498–520.

Hume, D., 1740. A Treatise of Human Nature. Penguin Books, reprinted in 1985, London.

Hummel, J.M., 2001. Supporting medical technology development with the analytic hierarchy process (pp. 1-186). Groningen.

International Monetary Fund, 2000. Offshore Financial Centers, IMF Background Paper,

Janský, P. 2013. Illicit Financial Flows and the 2013 Commitment to Development Index. CGD Policy Paper 034. Washington DC: Center for Global Development. http://www.cgdev.org/publication/illicit-financial-flows-and-2013-commitment-development-index

Jolliffe, I. T., 2002. Principal Component Analysis. Second ed. Springer Series in Statistics. New York: Springer-Verlag New York.

Justino, P., 2005. Empirical applications of multidimensional inequality analysis. Poverty Research Unit at Sussex Working paper.

Kaplowitz, M.D., Hadlock, T.D. and Levine, R., 2004. A comparison of web and mail survey response rates. Public opinion quarterly, 68(1), pp.94-101.

Kiesler, S. and Sproull, L.S., 1986. Response effects in the electronic survey. Public Opinion Quarterly, 50(3), pp.402-413.

Klasen, S., 2000. Measuring poverty and deprivation in South Africa. *Review of income and wealth*, *46*(1), pp.33-58.

Krosnick, J.A., 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied cognitive psychology*,*5*(3), pp.213-236.

Krosnick, J.A. and Presser, S., 2010. Question and questionnaire design.Handbook of survey research, 2, pp.263-314.

Krosnick, J.A., Holbrook, A.L., Berent, M.K., Carson, R.T., Hanemann, W.M., Kopp, R.J., Mitchell, R.C., Presser, S., Ruud, P.A., Smith, V.K. and Moody, W.R., 2002. The impact of" no opinion" response options on data quality: Non-attitude reduction or an invitation to satisfice?. Public Opinion Quarterly, 66(3), pp.371-403.

Lehmann, D.R. and Hulbert, J., 1972. Are three-point scales always good enough?. Journal of Marketing Research, 9(4), pp.444-446.

Lissitz, R.W. and Green, S.B., 1975. Effect of the number of scale points on reliability: A Monte Carlo approach. Journal of Applied Psychology, 60(1), p.10.

Lugo, M.A., 2007. Comparing multidimensional indices of inequality: methods and application. Research on Economic Inequality, 14, pp.213-236.

Lynn, P. and Jowell, R., 1996. How might opinion polls be improved?: the case for probability sampling. Journal of the Royal Statistical Society. Series A (Statistics in Society), pp.21-28..

Maasoumi, E. and Lugo, M.A., 2008. The information basis of multivariate poverty assessments. In Quantitative approaches to multidimensional poverty measurement (pp. 1-29). Palgrave Macmillan UK.Martin, W.S., 1978. Effects of scaling on the correlation coefficient: Additional considerations. Journal of marketing research, pp.304-308.

Mascherini, M. and Hoskins, B., 2008. Retrieving expert opinion on weights for the Active Citizenship Composite Indicator, European Commission, Joint Research Centre. Institute for Social Protection and Security of the Citizen.Mayer, S. E., and C. Jencks (1989): .Poverty and the distribution of material hard-ship,.Journal of Human Resources, 24(1), 88.114.

Moldan, B., Billharz, S. and Matravers, R., 1997. Sustainability indicators. A report on the Project on Indicators of Sustainable Development.

Murphy, R. 2008. Finding the screcy world: Rethinking the language of “offshore.” London: Tax Research LLP. http://www.taxresearch.org.uk/Documents/Finding.pdf ———. 2009. Where are the

world's secrecy jurisdictions? Tax Justice Network. Available online: http://www.secrecyjurisdictions.com/PDF/SJ_Mapping.pdf.

Murphy, R. and Vigueras, J.H., 2007. Tax Justice Network. Closing the floodgates–collecting tax to pay for development.

Myung, I.J., 2000. The importance of complexity in model selection. Journal of Mathematical Psychology 44, 190–204.

Nardo, M., Saisana, M., Saltelli, A. and Tarantola, S., 2005. Tools for composite indicators building. European Commission, EUR 21682 EN.Institute for the Protection and Security of the Citizen, JRC Ispra, Italy.

Oppenheim, A.N., 2000. Questionnaire design, interviewing and attitude measurement. Bloomsbury Publishing.

Oxfam, 2000. Tax Havens: releasing the hidden billions for poverty eradication.

Pearson, K., 1901. On Lines and Planes of Closest Fit to Systems of Points in Space. Philosophical Magazine, Series 6 , 2 (11), 559–572.

Popper, K.R., Kneale, W.C. and Ayer, A.J., 1948. Symposium: What Can Logic Do for Philosophy?. Proceedings of the Aristotelian Society, Supplementary Volumes, 22, pp.141-178.

Rabbani, S.J.R. and Rabbani, S.R., 1996. Decisions in transportation with the analytic hierarchy process. UFPB/CCT.a.

Ramsay, J.O., 1973. The effect of number of categories in rating scales on precision of estimation of scale values. Psychometrika, 38(4), pp.513-532

Saaty, R. W., 1987. The analytic hierarchy process- what it is and how it is used,. Mathematical Modelling, 9(3-5), 161.176.

Sen, A.K., 1997. On Economic Inequality, expanded edition with a substantial annexe by James E. Foster and Amartya Sen. *Oxford: Clarendon Press.(1979)," The Welfare Basis of Real Income Comparisons: A Survey," Journal of Economic Literature*, *17*(1), pp.1-45.

Scott, C., 1961. Research on mail surveys, Journal of the Royal Statistical Society. Series A (General), 124(2), pp. 143–205.

Sewell, M., 2008. Principal Components Analysis. Department of Computer Science University College London

Smallwood, A. D., 2014. An analysis of the efficacy of Tax Justice Network's methodology in constructing a secrecy index (Report prepared for: Cayman Finance). Retrieved from http://www.caymanfinances.com/wp-content/uploads/2014/08/TJNreportFinal.pdf

Srinivasan, V. and Basu, A.K., 1989. The metric quality of ordered categorical data. Marketing Science, 8(3), pp.205-230.

Stapleton, L. M. and Garrod, G. D., 2008. The Commitment to Development Index: An Information Theory approach. Ecological Economics, 66(2/3), 461–467. doi:10.1016/j.ecolecon.2007.10.009

Stapleton, L.M. and Garrod, G.D., 2007. Keeping things simple: why the Human Development Index should not diverge from its equal weights assumption. Social Indicators Research 84, 85–102.

Valencia, M. 2013. Storm survivors. Special reports of The Economist on offshore finance. The Economist, February 16. Available online: http://www.economist.com/news/special-report/ 21571549-offshore-financial-centres-have-taken-battering-recently-they-have-shownremarkable

Vaillancourt, P.M., 1973. Stability of children's survey responses. *Public opinion quarterly*, *37*(3), pp.373-387.

Witt, Karlan J. 1998. “Best Practices in Interviewing Via the Internet.” Proceedings of Sawtooth Software Conference, Sawtooth Software, Inc., Sequim, Washington, pp. 15-37.

Zoromé, Ahmed 2007: Concept of Offshore Financial Centers: In Search of an Operational Definition (IMF Working Paper), Washington D.C., in: http://www.imf.org/external/pubs/ft/wp/2007/wp0787.pdf

Appendix 1: The FSI 2015 rank

RANK	Jurisdiction	FSI Value	Secrecy Score	Global Scale Weight
1	Switzerland	1 466,1	73	5,625
2	Hong Kong	1 259,4	72	3,842
3	USA	1 254,8	60	19,603
4	Singapore	1 147,1	69	4,280
5	Cayman Islands	1 013,2	65	4,857
6	Luxembourg	817,0	55	11,630
7	Lebanon	760,2	79	0,377
8	Germany	701,9	56	6,026
9	Bahrain	471,4	74	0,164
10	UAE (Dubai)	440,8	77	0,085
11	Macao	420,2	70	0,188
12	Japan	418,4	58	1,062
13	Panama	415,7	72	0,132
14	Marshall Islands	405,6	79	0,053
15	United Kingdom	380,2	41	17,394
16	Jersey	354,0	65	0,216
17	Guernsey	339,4	64	0,231
18	Malaysia (Labuan)	338,7	75	0,050
19	Turkey	320,9	64	0,182
20	China	312,2	54	0,743
21	British Virgin Islands	307,7	60	0,281
22	Barbados	298,3	78	0,024
23	Mauritius	297,0	72	0,049
24	Austria	295,3	54	0,692
25	Bahamas	273,1	79	0,017
26	Brazil	263,7	52	0,678
27	Malta	260,9	50	0,990
28	Uruguay	255,6	71	0,037
29	Canada	251,8	46	1,785
30	Russia	243,3	54	0,397
31	France	241,9	43	3,104
32	Isle of Man	228,6	64	0,068
33	Liberia	218,2	83	0,006
34	Bermuda	217,7	66	0,042
35	Cyprus	213,9	50	0,518
36	Liechtenstein	202,4	76	0,010

37	Ireland	187,4	40	2,313
38	Belgium	181,2	41	1,863
39	Guatemala	177,2	76	0,007
40	Israel	173,8	53	0,166
41	Netherlands	168,4	48	0,322
42	Chile	166,7	54	0,120
43	Saudi Arabia	163,9	61	0,037
44	Australia	148,1	43	0,586
45	India	148,0	39	1,487
46	Philippines	146,1	63	0,020
47	Vanuatu	142,8	87	0,001
48	Ghana	139,2	67	0,010
49	Korea	124,3	44	0,302
50	US Virgin Islands	118,2	69	0,004
51	Samoa	117,5	86	0,001
52	Mexico	117,1	45	0,211
53	Norway	110,7	38	0,731
54	New Zealand	109,4	46	0,129
55	Gibraltar	109,3	67	0,005
56	Sweden	100,9	36	1,006
57	Aruba	99,5	68	0,003
58	Italy	98,7	35	1,218
59	Latvia	92,8	45	0,113
60	Belize	92,5	79	0,001
61	South Africa	90,9	42	0,203
62	Botswana	90,6	71	0,002
63	Anguilla	89,4	69	0,002
64	St Vincent & the Grenadines	79,7	78	0,000
65	Antigua & Barbuda	79,6	81	0,000
66	Spain	77,5	33	1,090
67	Costa Rica	74,9	55	0,010
68	Turks & Caicos Islands	72,5	71	0,001
69	St Kitts & Nevis	68,4	78	0,000
70	Curacao	67,8	68	0,001
71	Iceland	67,1	46	0,035
72	Seychelles	60,8	71	0,000
73	Slovakia	60,1	50	0,011
74	Macedonia	59,5	66	0,001
75	Poland	57,2	36	0,172
76	Monaco	53,7	74	0,000
77	Estonia	52,9	44	0,023
78	Portugal (Madeira)	52,5	39	0,063
79	St Lucia	51,7	83	0,000
80	Brunei Darussalam	47,4	83	0,000

81	Czech Republic	44,2	35	0,105
82	Grenada	42,2	76	0,000
83	Denmark	38,2	31	0,219
84	Hungary	37,3	36	0,052
85	Greece	37,2	36	0,046
86	San Marino	33,3	70	0,000
87	Andorra	27,3	77	0,000
88	Slovenia	22,5	34	0,019
89	Dominica	21,3	76	0,000
90	Finland	19,4	31	0,025
91	Cook Islands	17,8	76	0,000
92	Montserrat	10,9	67	0,000
NA	Taiwan	-	(67-75)	0,513
NA	Venezuela	-	(64-72)	0,230
NA	Dominican Republic	-	(65-73)	0,007
NA	Tanzania	-	(73-81)	0,006
NA	Montenegro	-	(60-68)	0,001
NA	Bolivia	-	(72-80)	0,001
NA	Paraguay	-	(75-83)	0,001
NA	Gambia	-	(73-81)	0,000
NA	Maldives	-	(76-84)	0,000
NA	Nauru	-	78,91	

Source: The Financial Secrecy Index 2015

Appendix 2: The list of fifteen KFSI

Number	KFSI
Knowledge of beneficial ownership	
1	**Banking Secrecy:** Does the jurisdiction have banking secrecy?
2	**Trust and Foundations Register:** Is there a public register of trusts/foundations, or are trusts/foundations prevented?
3	**Recorded Company Ownership:** Does the relevant authority obtain and keep updated details of the beneficial ownership of companies?
Key aspects of corporate transparency regulation	
4	**Public Company Ownership:** Does the relevant authority make details of ownership of companies available on public record online for free, or for less than US$10/€10?
5	**Public Company Accounts** Does the relevant authority require that company accounts are made available for inspection by anyone for free, or for a fee of less than US$10/€10?
6	**Country-by-country reporting**: Are all companies required to publish country-by-country financial reports?
Efficiency of tax and financial regulation	
7	**Fit for Information Exchange:** Are resident paying agents required to report to the domestic tax administration information on payments to non-residents?
8	**Efficiency of Tax Administration:** Does the tax administration use taxpayer identifiers for analysing information efficiently?
9	**Avoids Promoting Tax Evasion:** Does the jurisdiction grant unilateral tax credits for foreign tax payments?
10	**Harmful Legal Vehicles:** Does the jurisdiction allow cell companies and trusts with flee clauses?
International standards and cooperation	
11	**Anti-money Laundering**: Does the jurisdiction comply with the FATF?
12	**Automatic Information Exchange:** Does the jurisdiction participate fully in multilateral Automatic Information Exchange via the Common Reporting Standard?
13	**Bilateral Treaties:** Does the jurisdiction have at least 53 bilateral treaties providing for information exchange upon request, or is it part of the European Council/OECD tax convention?
14	**International Transparency Commitments:** Has the jurisdiction ratified the five most relevant international treaties relating to financial transparency?

15	**International Judicial Cooperation:** Does the jurisdiction cooperate with other states on money laundering and other criminal issues?

Source: The Financial Secrecy Index (2015)

Appendix 3: The survey

Notes:

- The numbering of the questions starts begin with 13 since it was 13th question on the FSI survey.
- The options depicted in question 1 were the same for all other questions.
- The underlined text at the end of each question was a hyperlink that lead to webpage containing details about the particular component.

Weighing of the 15 Key Financial Secrecy Indicators (KFSIs):

Currently, the weights of all KFSI indicators are set to be equal (every one of 15 indicators has its weight equal to 1).

Jurisdictions are awarded points for each indicator according to their transparency. If a jurisdiction is awarded a point, it suggests that it is transparent. If a jurisdictions obtains 100% of points then it is 100% transparent in terms of these indicators. More details can be found in the Methodology.

Please choose for each of the indicators, whether you think its weight should remain the same (equal to other indicators),or (significantly) lower/higher than other indicators.

Alternatively, you can assess concrete weights to each indicator. The sum of the weights can be random, e.g. if you think that the first indicator (Banking Secrecy) should be weighted twice as much as the second indicator (Trust and Foundation Register) you can choose the field ‘other’ and write “2” for the weights of the first one and respectively “1” for the second.

If you think that a particular indicator should be omitted, please enter 0 under the field "Other".

13. KFSI 1: Banking secrecy

The indicator assesses whether a jurisdiction provides banking secrecy. To obtain full credit, the jurisdiction must ensure that banking data exists, and that it has effective access to this data. More details can be found here.

Significantly Lower	Lower	The same	Higher	Significantly higher	Other	I don't know
○	○	○	○	○	○	○

14. KFSI 2: Trust and Foundation Register

This indicator reveals whether a jurisdiction has a central register of trusts. All trusts and private foundations formed and administered in a jurisdiction must be required to register with a central agency. More details can be found here.

15. KFSI 3: Recorded Company Ownership

This indicator assesses whether a jurisdiction requires all types of companies to submit beneficial ownership information to a governmental authority, and whether it requires this information to be updated, regardless of whether or not this information is made available on public record. More details can be found here.

16. KFSI 4: Public Company Ownership

This indicator considers whether a jurisdiction requires all available types of company with limited liability to publish updated beneficial ownership or legal ownership information on public records accessible via the internet. More details can be found here.

17. KFSI 5: Public Company Accounts

This indicator shows whether a jurisdiction requires all types of companies with limited liability to file their annual accounts and makes them readily accessible online via the internet at a maximum cost of US$ 10 or € 10. More details can be found here.

18. KFSI 6: Country by Country Reporting

This indicator measures whether the companies listed on the stock exchanges or incorporated in a given jurisdiction are required to publish worldwide financial reporting data on a CBCR (country by country reporting) basis. More details can be found here.

19. KFSI 7: Fit for Information Exchange

This indicator asks whether resident paying agents (e.g. stock companies and financial institutions) are required to report to the domestic tax administration information on all payments (of dividends and interest) to all non-residents. More details can be found here.

20. KFSI 8: Efficiency of Tax Administration

This indicator shows whether the tax administration of a given jurisdiction uses taxpayer identifiers for efficiently analysing information, and whether the tax administration has a dedicated unit for large taxpayers. More details can be found here.

21. KFSI 9: Avoids Promoting Tax Evasion

This indicator shows whether a jurisdiction grants unilateral tax credits for foreign tax paid on certain foreign capital income when remitted home. The types of capital income included are interest and dividend payments. More details can be found here.

22. KFSI 10: Harmful Legal Vehicles

This indicator has two components. It shows whether a jurisdiction allows the creation of “protected cell companies” (also known as “incorporated cell company” or “segregated account company”). Additionally, it measures whether the administration of trusts with flee clauses is prohibited. More details can be found here.

23. KFSI 11: Anti Money Laundering

This indicator examines the extent to which the anti-money laundering regime of a jurisdiction is considered effective by the Financial Action Task Force (FATF), the international body dedicated to counter money laundering. More details can be found here.

24. KFSI 12: Automatic Information Exchange

This indicator registers whether a jurisdiction participates in multilateral automatic information exchange on tax matters. Participation in the European Savings Tax Directive (EUSTD) is taken as a proxy for this indicator. More details can be found here.

25. KFSI 13: Bilateral Treaties

This indicator examines the extent to which a jurisdiction has signed and ratified bilateral treaties conforming to the 'upon request' information exchange standard developed by the OECD and the Global Forum with 46 other countries, and/or whether the jurisdiction has signed and ratified the Amended Council of Europe / OECD Convention on Mutual Administrative Assistance in Tax Matters. More details can be found here.

26. KFSI 14: International Transparency Commitments

This indicator measures the extent to which a jurisdiction has entered into international transparency commitments. We have checked whether a jurisdiction is party to five different international conventions. A credit of 0.2 points is awarded for each of the specified conventions. More details can be found here.

27. KFSI 15: International Judicial Cooperation

This indicator measures the degree to which a jurisdiction engages in international judicial cooperation on money laundering and other criminal issues. We use the degree of compliance with the Financial Action Task Force recommendations 36 through 40 as the appropriate measure. More details can be found here.

Appendix 4: Nationalities of respondents

Jurisdiction	# of responses	Jurisdiction	# of responses
United Kingdom	12	Italy	1
Germany	7	RSA	1
Switzerland	5	Japan	1
Belgium	4	Canada	1
USA	4	Argentina	1
Denmark	3	Netherlands	1
Spain	3	Latvia	1
Austria	3	Czech republic	1
Australia	3	Liberia	1
Portugal	2	Brazil	1
Norway	2	Liechtenstein	1
Ecuador	2	Seychelles	1
Netherlands	2	Luxembourg	1
France	2	Burundi	1
Kenya	2	Mauritius	1
Poland	1	Gambia	1
Cayman Islands	1	Mexico	1
Slovenia	1	Guatemala	1
India	1	Cyprus	1
Nicaragua	1		

Source: survey results

Appendix 5: Testing new weights against equal weights using t-test

I tested hypothesis that new weight for $KFSI_i$ is equal to 1 (equal weight) using t-test in Stata for each component. Therefore I performed 15 t-tests. The following table presents the aggregated results.

KFSI	mean	SE	SD	t	p-value
KFSI1	1,054	0,029	0,254	1,838	0,070
KFSI2	1,016	0,033	0,276	0,479	0,633
KFSI3	1,042	0,029	0,247	1,459	0,149
KFSI4	0,995	0,029	0,238	-0,161	0,873
KFSI5	0,954	0,028	0,239	-1,633	0,107
KFSI6	1,038	0,034	0,298	1,109	0,271
KFSI7	1,030	0,027	0,232	1,107	0,272
KFSI8	0,959	0,033	0,283	-1,249	0,216
KFSI9	0,944	0,036	0,297	-1,568	0,122
KFSI10	1,009	0,030	0,253	0,301	0,764
KFSI11	1,022	0,026	0,223	0,816	0,417
KFSI12	1,036	0,030	0,253	1,211	0,230
KFSI13	0,896*	0,032	0,269	-3,239	0,002
KFSI14	0,974	0,032	0,270	-0,804	0,424
KFSI15	1,030	0,028	0,242	1,060	0,293

* Significance of difference from equal weights at the 95 % confidence level

Source: own estimation; Stata software output