**A study on how the number of officially recognized languages within a country affect its economics freedom**

Module 7 – Capstone – Critical Thinking #7

Witt Shoaf

Colorado State University Global

MIS581-1 Capstone: Business Intelligence and Data Analytics

Dr. Steve Chung

5/6/2023

**Tables of Contents**

Abstract – 3

Introduction – 4

Problem/Purpose Statement – 5

Objectives – 5

Overview of Study – 6

Research Question and Hypotheses – 6

Literature Review – 7

Dataset Description – 9

Analytical Toolset – 9

Data Variables – 10

Methodology – 11

Limitations – 12

Ethical Concerns – 12

Findings – 13

Conclusion – 15

Recommendations – 15

References – 16

**Abstract**

All people use language to function within human society. Language is one of the most important differentiating features of our species. As humanity matured and grew more capable of self-organizing and developing structured communities, language also played a vital role in how our economies/market were shaped. The inability to conceptualize specific types of trades, economic concepts, or consequences over the long term are all associated with the language to some degree. This study seeks to understand a specific aspect of languages’ impact on economies, namely the economic freedom of citizens within their country. This analysis is done using linear regression on a dataset comprised of different sources which house two main variables. The results find that language does impact economic freedom of a country as a whole.

**Introduction**

For this paper, the intention is to investigate the relationship between spoken languages officially recognized and the economic freedom of the citizens living there. More specifically this research sets out to see if a country having more than one official language impacts the economic freedom rating, they receive from the Fraser Institute. Economic freedom is described as, “the fundamental right of every human to control his or her own labor and property” (Hertiage Foundation, 2023). This study is important for both local and federal level policy makers along with business owners. Throughout humans’ recent history a great competition over which economic model would become supreme was waged with capitalism/socialism being the winner (Cliffnotes, 2023). Understanding this to be how the current state of economic affairs is arranged for most modern countries. This study will provide statistical evidence that recognition of more than one official language within a country either impacts or doesn’t impact the country’s economic freedom. Armed with this knowledge policy and business strategy can be better determined for individuals/leaders on how to best shape their plans.

For the purpose of this research, the Economic Freedom Rankings dataset from Fraser Institute was selected as it has data on all countries dating back to 1970 (Source URL: <https://www.fraserinstitute.org/economic-freedom/dataset?geozone=world&page=dataset&min-year=2&max-year=0&filter=0>). With this amount of historical detail, we believe that sufficient evidence can be found to either prove or disprove the hypothesis which is stated later. The Fraser Institute provides a more substantial description of economic freedom which adds integrity to the analysis that the data being used is valuable and trusted by other researchers. A small exert detailing how economic freedom is defined within the actual data is:

“In an economically free society, the primary role of government is to protect individuals and their property from aggression by others. The EFW index is designed to measure the extent to which the institutions and policies of a nation are consistent with this protective function and the freedom of individuals to make their own economic decisions. Put another way, the EFW measure is an effort to identify how closely the institutions and policies of a country correspond with a limited government ideal, where the government protects property rights and arranges for the provision of a limited set of “public goods” such as national defense and access to money of sound value, but little beyond these core functions. In order to receive a high EFW rating, a country must provide secure protection of privately owned property, a legal system that treats all equally, even-handed enforcement of contracts, and a stable monetary environment. It also must keep taxes low, refrain from creating barriers to both domestic and international trade, and rely more fully on markets rather than government spending and regulation to allocate goods and resources.”

(Fraser Institute, 2022)

This description of economic freedom by the actual dataset owners is important to keep in mind. The overall goal of this analysis is to see if a country having 2 or more official languages impacts the value represented by this description.

**Problem/Purpose Statement**

As spoken/written language is the humanities’ primary means of communication it goes without saying that understanding its impact on our daily lives will be beneficial. Various studies on the beneficial impact of being multilingual have not provided definitive proof of any measurable benefit it provides to an individual. Understanding that these studies were done by looking only at the individual we have decided to investigate if a country being multilingual impacts, it in a substantial way. While the effects of multilingualism might be hard to discern on an individual basis, looking more wholistically at a total country in comparison to its peers might shed light on statistically significant benefits that occur from multilingualism. The overarching problem statement is this, does multilingualism impact economic freedom.

**Objectives**

With the findings of this we are hopeful that new and ultimately fruitful discussion can occur about the benefits or non-benefits of multilingualism as a public policy. The authors are in no way advocating for either and understand the personal benefits that multilingualism might provide an individual. This study is aimed at providing evidence-based analysis by which policy makers and business owners can consider future strategic moves to benefit their constituents and stakeholders. Since no company and or country is completely independent, understanding how internal language impacts the livelihoods of a country’s people can help leaders shape their decisions to maximize returns on investments.

Additionally, the goal of this analysis is to provide a steppingstone for future research in this specific field. For a long time, economic thinkers considered humans as rational beings which made only rational choices. With the introduction of behavioral economics, a new view of people being individualist and irrational was adopted. With this consideration in mind this research is being conducted to see the impacts on a macro scale of specific policy and the overall outcome it produces for individuals within the system. To that end this research will hopefully provide aid in other research within the subjects of behavioral economics, macroeconomics, government science, and other fields relating to humanities.

**Overview of Study**

This study will produce through qualitative analysis an understanding of if a relationship exists between a nations official language count and its economic freedom. Using publicly available data from the Fraser Institute, Wikipedia, and the University of Ottawa the final dataset has been put together. Regression analysis will be used to determine the relationship through predictive analysis. The goal as previously stated is to add to the growing knowledge base concerning this approach of viewing economic freedom.

**Research Question and Hypotheses**

The research question is this, does a country being officially multilingual benefit it in the form of higher-than-average economic freedom for its citizens. The hypotheses developed to test this question are:

H0: Official multilingualism will not produce any difference in economic freedom on average as measured by the beta of the linear regression model being 0.

H1: Official multilingualism will produce a difference in economic freedom on average as measured by the beta of the linear regression model being anything but 0.

Using this hypothesis test will allow the research to provide a concrete understanding of the impact that official multilingualism has on countries through the form of its economic freedom ratings. Developing and refining the research question was done in part through the teaching of Siah Ang (Ang, 2014).

**Literature Review**

Several different sources were reviewed concerning how language impacts the economic strength/freedom of a country. Most of the research analyzed how language might impact an individual’s behavior, how language structures impact saving and spending habits, and how specific languages impact economic growth. All of these analyzes help shed light on different ways to consider tackling the central problem of this research paper, does a country being multi-language impact the economic freedom score it receives.

Understanding the impact of language on economic freedom requires thinking about language as an impactful variable on economies which can either hinder or improve economic growth of a country. In a paper considering the impact of the English being the medium of instruction for students the case is plainly stated that the use of English can be considered beneficial for future economic growth. When considering how a country might view adopting a multi-language policy the impact of adopting English as the second standard language could have considerable beneficial impacts. Vaishnav sums this up with a simple statement concerning how one’s primary language can impact one’s business ability, “It’s easy to take for granted being born in a country where people speak the *lingua franca* of global business (English)” (Vaishnav, 2020). If one can speak English, their ability to conduct business globally and thus positively impact their economic potential would be higher than someone who is unable to speak English. Along this same vein of research Lee found that there is, “evidence of positive correlation between initial English proficiency and economic growth only for the countries in the Asia and Europe” (Lee, 2012). While both studies only look at the impact of English on economic growth for countries (which this research takes to positively impact economic freedom), this lends some credibility to the idea that being multi-lingual as a country could positively impact economic freedom of the individual. Thinking more broadly there would be other benefits to speaking another language besides English, although these benefits could be less (this is not proven or understood and should not be taken as a fact).

Other literature worked along the idea that language might be a mechanic by which humans evaluate the trustworthiness of others subconsciously. Chong(2012) considered in their *Language nuances, trust and economic growth* the impact of language on helping people develop trust in a business setting. Chong considers how similarity in language typically signifies similarity in shared values, culture, beliefs, and customs (Chong, 2010). This human characteristic of seeking people from one’s own culture does lend itself to the idea that being multilingual could help a country establish stronger economic bonds with other countries through the individual bonds of people speaking the same language.

The last piece of literature considered concerns how language impacts an individual’s behavior. The specifics here are far reaching on their impacts to an individual and a country. Chen (2013) in their paper explains how certain languages describe the future/past very specifically and intentionally within the language thereby separating it from the present (in English the future is something that *will* happen, and the past is something that has *happened* as an example). Other languages, they go on to describe, blend all time flows together when speaking and let the context of the situation indicate to the listener what the speaker means. Their findings, while not conclusive/exhaustive indicate that languages with no clear separation of future/past from the present tend to have their speakers act in more future-oriented behavior such as saving for retirement more. Chen’s explanation for this is that without the separation of time being spoken by the person they might view the future/past/present as one and the same and more easily see the value of saving for tomorrow because it is the same as today. While Chen’s research is extremely interesting on the impact of language for individuals, they did not conduct any further analysis on the impact that being multilingual in a language the blurs time has on an individual’s savings.

The current academic literature has been circling the general concept that language does in fact impact the economic growth/strength of a country through differ means of analysis and research. Understanding that there is evidence of impacts through language on human capital, smoothing of translation and trust between business partners, and an individuals’ behavior, provide evidence that language does in fact contribute to a country’s economy. While the research pursued in this paper does not seek to definitively answer how languages impact economic growth/strength is does attempt to add to the available literature on languages impact on economies.

**Dataset Description**

For this assignment, the Economic Freedom Master Index is be used as the primary dataset. It lists the year, 165 individual countries, their economic freedom summary index score, and the quartile (this data runs from 2020 to 1970). The authors in addition to this dataset will add official language counts to the respective countries using online resources. This additional row of data will be either a 1 (single official language) or a 2 (multiple official languages). This distinction is because many countries have more than 2 official languages and the exact count is not being considered, but the actual occurrence of more than one official language. The sourcing of this data will occur in whatever online resource is available and the running assumption is that all countries with more than one official language will have it applied from the beginning of the dataset to the end (specific years will not be considered). The logic for this is that most countries are older than 1970 and will have had their other official languages in use for decades if not centuries prior to the law/policy catching up to recognize it officially.

**Analytical Toolsets**

Two software applications will be used to conduct this analysis. The first tool will be the SAS Academics on Demand version of SAS. This tool will work on the major analytical heavy lifting required to complete the project. The goal of this research project is to analyze the relationship between the occurrence of official languages within a country and its economic freedom. Most of the analysis will center around regression. SAS has a wealth of statistical analysis tools and can even conduct correlation without issue (SAS, 2023). The date will be cleaned and prepared before being loaded into the software to be analyzed. Its output consisting of information tables will be utilized when writing about the hypothesis and the outcome of the research. Additionally statistical description analysis might be conducted using SAS for the first pass as it will help identify any unknown issues that might occur within the data itself.

The second tool will be Python. Python will specifically be used to create visuals of the output and the data itself for the research paper. This is because Python allows a great degree of freedom in making visuals which will help with the story telling of the analysis and its outcome (Elliot, 2022). For the analysis python was not chosen as the SAS software already has predefined analytics tools which allow for rapid and iterative analysis to be conducted.

Lastly a special word is required concerning that excel will be used on the first pass of the dataset. Excel will host the base data and allow the authors to add and make crucial fixes within the dataset as needed. As a part of the analysis other data might be added beyond the official language count such as average temperate for the year, geographic location, common currency, and or average GDP per capita. These other data pieces will help the authors understand more contextual information surrounding the effect of official language, but most likely will not appear in the final paper unless they help tell the story of the findings.

**Data Variables**

The dataset is rather simple which is why additional data will be added manually. The different variables including the context adding variables mentioned above are.

1. Years: this variable is a date (dates are in year form and range from 1970 to 2020)
2. Countries: this variable is a string (additional analysis will need to be consider based on new or lost countries and any potential name changes)
3. Economic Freedom Summary Index: this value is an integer (made up several other values which will be explained within the final research paper)
4. Languages Spoken: this variable is an integer (as mentioned above 1 will represent 1 official language while 2 will represent any value above 1)
5. Average Temperate per Year: this variable will be an integer.
6. GDP per Capita: this variable will be an integer.

Using these variables, the analysis will be conducted. Basically, the variables are integers as this is a quantitative analysis which requires the ability to statistically analyze the data (Ali, 2021).

Figure #1

Data Dictionary Visual

A picture containing text, screenshot, font, line

Description automatically generated

Source: Produced by the Authors within Python

**Methodology**

The research question will be answered by using linear regression within SAS. Finding statistical significance through the p-values will be the primary determining factor. Our process will consist of developing the entire dataset and ensuring its completeness as best as possible. Then after developing descriptive statistical views plugging the variables into the linear regression model as such: *dependent variable:* freedom index by year per each country, *independent variable:* number of official languages spoken and year of data. This isolated test will provide evidence of if having multiple languages does impact economic freedom. The results will be interpreted by 1) evaluating the respective p-value for its significance and 2) considering the resulting slope equation and intercept value as being negative or possible (if multi-languages predict a positive outcome we will take that as the answer) (Sarstedt, 2014). If additional tests are needed to help solidify the answer, they will be considered at the point with either a neural network and or logistic regression being the selected backups. Special focus will be kept with linear regression as the researchers feel it is the easiest to understand and explain the logic of anyone new to statistical analysis.

**Limitations**

The main body of the data used for the analysis relates to the Economic Freedom index value provided by the Fraser Institute database. This dataset, while expansive and publicly available, does have missing index data due to data gathering limitations of the Fraser Institute. Data began being gathered in 1970 with only 84 of the 165 total countries have values available for them in the beginning. Additionally, for the years preceding 2000 the data is only available every 5 years. This gap in data coverage will limit the longitudinal value of the research.

The other main aspect of the data, the independent variable (or official language count) also has a major limitation. The researchers of the paper have opted to consider multiple official languages for a country being “permanent” due to data limitations on when official languages were enacted. The justification for taking this approach has been described previosuly.

**Ethical Concerns**

The authors of this research paper do not foresee any ethically related issues with the data selected besides individuals taking some perverted sense of superiority in economic terms based on individual languages. This we hope will not occur and recognize is entirely outside of the control of the researchers to impact. The paper will be written neutrally to minimize this occurrence. As for the source of the data itself, no ethical issues have been found during the researcher’s investigation in either its sourcing or its ability to impact individuals.

**Findings**

The findings of the analysis show that the null hypothesis is disproven. The confidence of these findings is validated by the p-values both the intercept and the language count being well below the threshold for analysis (0.05). The null hypothesis predicted that language counts would have no impact on the economic freedom of a country which is disproven by the slope value being nonzero within the regression model. As can be seen in Figure #2 the regression model predicts that multiple languages will cause an impact on the average economic freedom (although negatively).

Figure #2

Regression Results from SAS

Table

Description automatically generated

Source. SAS Studio for Academics used by the researchers to conduct their analysis.

The input data took the average economic freedom for each year of data and compared each individual country against it. A value of 1 was given to countries with above average economic freedom while a value of 2 was given to countries with below average economic freedoms. Language was similarly simplified for the analysis by giving countries a value of 1 if they use a single language and 2 if they use any number of languages greater than 1. The resulting equation this regression model predicts is Economic Freedom Average Value = 1.40451 + (language count) \* 0.05363. The result shows that with the inclusion of multiple languages (language count = 2) the country would have an Economic Freedom Average Value of 1.51177 (1.40451+(2)\*0.05363) thereby signifying that the country’s economic freedom is lower than average. Figure #3 which was taken from the same SAS regression model shows how simplified that data was by using (1 or 2) when showing values for the above/below average freedom index and language variables.

Figure #3

Scatterplot of Results

A picture containing text, screenshot, display, number

Description automatically generated

Source. SAS Studio for Academics used by the researchers to conduct their analysis.

**Conclusions**

This research set out to determine if a country being multi-lingual was beneficial for its citizens economic freedom. The findings suggest that multiple languages in fact are related to lesser economic freedom on average for the country’s citizens. The more concrete finding is that language does impact economic freedom.

**Recommendations**

While the findings of this research go against commonly held beliefs that being multi-lingual is beneficial for individuals and their communities, the researchers would like to caution against any hasty decisions. The data itself was taken at a very high level and oversimplified to allow for the research being completed in the allotted time. The researchers noted while compiling the data that certain regions of the world appeared to have higher amounts of official languages than others (this most likely being related to colonialism and the arbitrary country borders it created). While the researchers do not wish to characterize these areas as being more economically challenged than other areas within the study it would stand that this might be the case. The results then are more than likely too generalized, but still hold value in having looked at the research material and identifying an approach to its study. Additional research should be conducted on a more regionalized basis that does not compare wealthier areas of the world with poorer uniformly. Instead, any next research in this subject should seek to establish groups of countries with similar economic conditions and then compare them to their peers through the method this research established.

For the policy/business leaders this study should serve more as a warning that being over generalized in analysis will lead to incomplete understandings of the cause and effects of individual variables.

**References:**

Ang, S. (2014). The research question. SAGE Publications Ltd, Retrieved from <https://doi.org/10.4135/9781473909694>

Chen, M. (2013). The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets. American Economic Review. Retrieved from <https://web-s-ebscohost-com.csuglobal.idm.oclc.org/bsi/pdfviewer/pdfviewer?vid=0&sid=bbe91963-12c5-472c-aa15-741434813104%40redis>

Chong, A. (2010). Language nuances, trust and economic growth. Public Choice. Retrieved from <https://web-p-ebscohost-com.csuglobal.idm.oclc.org/bsi/pdfviewer/pdfviewer?vid=0&sid=f9af9ea6-37ad-4cbd-9ca9-51d346f9207f%40redis>

CliffNotes. (2023). Predominant Economics Systems. CliffNotes. Retrieved from <https://www.cliffsnotes.com/study-guides/sociology/economics-and-politics/predominant-economic-systems>

Fraser Institute. (2022). Economic Freedom Basics. Fraser Institute. Retrieved from <https://www.fraserinstitute.org/economic-freedom/economic-freedom-basics>

Heritage Foundation. (2023). 2023 Index of Economic Freedom. Heritage Foundation. Retrieved from <https://www.heritage.org/index/about#:~:text=What%20is%20economic%20freedom%3F,in%20any%20way%20they%20please>.

Lee, C. (2012). English Language and Economic Growth: Cross-Country Empirical Evidence. Journal of Economic & Social Studies. Retrieved from <https://web-s-ebscohost-com.csuglobal.idm.oclc.org/bsi/pdfviewer/pdfviewer?vid=0&sid=00123e9c-7152-485d-b845-c90f340e231d%40redis>

Sarstedt, M. (2014) Regression Analysis. University of Melbourne. Retrieved from <https://www.researchgate.net/publication/300403700_Regression_Analysis>

SAS. (2023). SAS OnDemand for Academics Features. SAS. Retrieved from [https://www.sas.com/en\_us/software/on-demand-for-academics/features-list.html#](https://www.sas.com/en_us/software/on-demand-for-academics/features-list.html)

Vaishnav, D. (2020). Impact of medium of instruction in education on economic growth and development. Theoretical and Applied Economics. Retrieved from <https://web-p-ebscohost-com.csuglobal.idm.oclc.org/bsi/pdfviewer/pdfviewer?vid=0&sid=4ef86a46-8ccd-4847-906d-ce62be63b997%40redis>

Weber, S. (2020). The Economics of Language. Journal of Economic Literature. Retrieved from <https://www.aeaweb.org/articles?id=10.1257/jel.20191316>