STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
659 Degree project in economics
Spring 2016

# THE MACROECONOMICS OF LANGUAGE:
## A Study on Linguistic Diversity and National Productivity

Marielle Nyberg (22832) and Alexander Kidelius (23071)

Abstract: Using a cross-section of linguistic and economic data on the 34 member states of the OECD, this paper studies the effect of linguistic diversity on economic productivity in economically developed countries. Linguistic diversity is measured through five different indices in order to find the measure that best explains the variance in Total Factor Productivity. We find, in contrast to existing research, that all measures surveyed indicate a positive relationship between linguistic diversity and TFP, and that the Esteban-Ray index provides the most statistically significant result wherein an increase of one percentage in ER corresponds to an increase in TFP by 1.9 %. A casual relationship is investigated but cannot be established, as the positive coefficient may be the result of reverse causality wherein productivity increases linguistic diversity.

Keywords: Economic Anthropology, Communication barriers, Ethnolinguistic groups, Polarization, Fractionalization

JEL: C31, D70, H5, O5, Z10

Supervisor: Karl Wärneryd
Date submitted: May 16, 2016
Date examined: June 2, 2016
Discussant: Myrenfors, S. and Shams, A.
Examiner: Maria Perotta Berlin

# ACKNOWLEDGEMENTS

BS"D

We would like to extend our thanks to our supervisor, Karl Wärneryd, for his constructive comments and assistance throughout the writing of this thesis.

Further, we owe gratitude to Professor Paul Segerstrom at the Department of Economics of the Stockholm School of Economics for the input and support he provided.

In addition, we seek to express our appreciation for the speedily and most helpful assistance with obtaining data that we received from Professors Romain Wacziarg and Enrico Spolaore at the UCLA Anderson School of Management and Tufts University, respectively.

# TABLE OF CONTENTS

# 1.0 Introduction

Consider the case of a society composed of different groups who are unable to effectively communicate with one another. Most likely, difficulties will arise in collaborating, every step of the way from manually interacting, to agreeing on goals and the methods for achieving these. It is less than a stretch to from there develop the belief that poor communication prospects could hamper a society's economic development and welfare. Consider then, instead, the case of a wildly prosperous society that aptly realizes and utilizes its citizens' potential. Such a country would likely tend to attract migration of individuals from varied backgrounds and tongues, thus fostering linguistic diversity as an indirect result of its economic success. And perhaps it is even so that the variation in language could itself enable the presence of a greater diversity in ideas, making the effect thereof the very opposite of the first hypothetical scenario.

This prompts the question: *what is the true relationship between economic success and linguistic diversity in developed countries?* If we were able to arrive at whether linguistic diversity has a positive or negative economic impact, this could potentially assist in making efficient language policy decisions, for instance in choosing whether to standardize language use or to encourage the retention of a larger multiplicity of tongues.

The current state of knowledge on this topic suggests that linguistic diversity negatively affects economically-related variables such as growth, institutions, and political stability (Alesina and Ferrara 2005, Esteban and Ray 1994). We, however, will herein propose that this is not necessarily universally true. Instead, we suggest that linguistic diversity can be positively correlated with indicators of economic success in developed countries. This is relatedly consistent with Florida and Gates (2001), Collier (2001), and Florida (2002) who also found examples of diversity in a positive way affecting economic variables.

As we are interested in the communicative mechanism of how language affects economic success, we would like to examine an economic dependent variable that is related to these mechanisms. For this reason, we will focus on a productivity measure—Total Factor Productivity—rather than output measures such as GDP. The reasoning therefor will be explained in greater detail below.

We use linguistic diversity to measure the inter-societal communicative ability of a country and the level of antagonism that this might create. That is, how diversity generates feelings of identification and alienation within and between groups, respectively. The measures we consider have already been used extensively in existing research and they can generally be grouped into fractionalization measures—expressing the degree to which a population is fractured into various groups—and polarization measures, which explain occurrences with focus on disparities in group sizes. Some of these measures also account for lexical distances between languages, while others do not.

From this, we seek to investigate whether one of these measures can be proven to explain a greater share of countries' variation in productivity than the others, in order to be able to obtain the optimal choice of linguistic diversity measure for investigating the relationship between countries' levels of total factor productivity and language diversity.

This paper will be structured as follows. First, we will review existing literature and research relating to language economics and diversity, which will serve as a base for our hypotheses presented in section three. Section four deals with methodology as well as data and measurement issues. The empirical results will be presented and discussed in section five, and in section six the problem of reversed causality is discussed. Section seven concludes and offers concluding statements on the hitherto presented material.

## 2.0 Literature review

Below is a summary of some of the key findings on the economics of language. First is a short description of the relationship between the benefits and costs of linguistic diversity. This is followed by a review of literature explaining communication issues and their implication on coordination and efficiency. Last is a summary of the key findings on linguistic diversity on productivity and growth.

### 2.1 Previous research—Languages and economic success

Economists have long attempted to answer the question of what determines economic success, and in doing so have often been focusing on human and physical capital, technology, and institutional origins. With the introduction of the concept of linguistic diversity, a related field of study emerged which examined linguistic diversity as an important determinant for economic performance (Grafton, Knowles and Owen, 2004, Desmet, Weber and Ortuño-Ortín, 2009, Alesina et al., 2003, Alesina and Ferrara, 2005, Grafton, Kompas and Owen, 2007).

Ashraf and Galor (2013) argue that there are two separate forces of diversity that have the opposite effect on performance and growth. On the one hand, diversity can be positive for technological development as it allows for complementary ideas between people with different backgrounds, which could serve to boost innovation—pushing out the production possibility frontier. On the other hand, higher levels of diversity may instead make it harder for countries to operate at their production possibility frontier as it could reduce trust and coordination, as well as potentially the exchange of ideas by making communication more difficult.

Consider, for instance, the game-theoretical classic *The Stag Hunt* game. In this game, we see that two players unable to communicate with one another tend to arrive at the lower right quadrant {Rabbit, Rabbit} achieving the equilibrium of {1,1}. This is because the risk of obtaining zero, in the event that oneself chooses *Stag* and the other person plays *Rabbit* giving {0,2} or {2,0}, are too large if one cannot be assured that the other player will play *Stag*. Therefore, the probability is large that two players—in the absence of communicative prospects—will play the risk-dominant Nash equilibrium in the lower right quadrant, rather than the pay-off dominant such in the upper left.

**Figure 1. The Stag Hunt game**

| | | **Player 2** | |
|---|---|---|---|
| | | Stag | Rabbit |
| **Player 1** | Stag | 3,3 | 0,2 |
| | Rabbit | 2,0 | 1,1 |

(1.0)

The greater their abilities to communicate, the greater is the probability that the two players will be able to coordinate on the payoff-optimal equilibrium of {3,3}. This theoretical example can be extrapolated onto the real-life context of an entire country – because any economy ultimately consists of a myriad of small transactions and small-scale coordination games, to express it in a simplified way. And though the real-life situations may not always be as stark as in the above game, the basic mechanisms thereof ultimately add up on the macro scale as well.

These two forces associated with diversity on a linguistic plane must be balanced so that the potential positive effects of innovation are not outweighed by the potential negative implications of reduced communicative prospects. Ashraf and Galor (2013) found that the countries that performed the best were those with intermediate levels of diversity. This hypothesis was also tested and supported by Desmet, Ortuño-Ortín and Wacziarg (2014).

### 2.1.1 Literature on communication

Language is one of the most fundamental elements of communication. It can also, according to some economists, influence thought and behavior to the degree that differences between languages create differences in how their speakers think and behave (Gay et al. 2013). For example, Gay et al. arranged languages according to gender intensity—id est to what degree the nouns in a language are grouped into gendered classes. They found that women who speak a language that marks gender more intensively are less likely to participate in economic and political activities.

Santacreu-Vasut et al. (2014) add that countries with dominant languages with high gender marking tend to have less female presence on boards and committees than countries with lower such marking. This is one source of communication inefficiency. If language can determine thought and culture, then understanding and trust between individuals that speak different languages will be impeded. Another example where language affects thought is proposed by Chen (2013). Chen studied the effect of time-referencing on behavior and found that speakers of languages that mark tense more strongly will engage less in long-term behavior such as actions towards health and savings. This can cause coordination problems between individuals that speak native languages that have strong and weak time-marking, respectively.

Similarly to the language and thought relation, there are also studies on the emotional implications of language. Languages possess different levels of nuances and precision regarding the expression and description of emotions. As thoughts are not always rational but emotional, the way and emotions are expressed in language will affect the way individuals think and behave. Language is so closely intertwined with culture that not everything can be fully translated. This means that feelings transferred using words can create recognition, but it can also alienate non-native speakers of the language employed to do so (Bond and Ginsburgh, 2016).

### 2.1.2 Literature on productivity and growth

The majority of studies on the economics of language do not focus on communication between individuals. They are on the macroeconomic level intended to explain variation between countries. Many of these focus on *how* to measure linguistic diversity, but they also present hypotheses and empirical results on what effect these have on a number of economic outcomes. The consensus is clear—the majority present linguistic diversity as an impediment to economic growth (Grafton, Knowles and Owen 2004, Grafton, Kompas and Owen 2007, Alesina and Ferrara 2005) and institutional efficiency and corruption (Mauro, 1995), as well as a source of ethnic conflict and civil unrest (Esteban and Ray, 1994, Reynal-Querol, 2002).

In order to define the relationship between linguistic heterogeneity and economic performance, researchers have developed a number of different indices of linguistic diversity. These indices can be sorted into two classes: fractionalization indices and polarization (or disenfranchisement) indices. Within both of these classes are subgroups of measures that either ignore linguistic distances or that include them.

The most widely used index for ethnolinguistic fractionalization is the ELF index (Atlas Narodov Mira 1964, Easterly and Levine 1997, Alesina et al. 1999). It was first developed by a group of Soviet researchers (Atlas Narodov Mira 1964) and has the most extensive set of ethnolinguistic data to this day. Easterly and Levine (1997) used the ELF index to explain cross-country differences in growth rates as well as public policies and political stability in Sub-Saharan Africa. They found that most of the characteristics of Africa's poor growth, including high government deficits and underdeveloped financial systems, were closely associated with high ethnolinguistic fractionalization. They also broadened the scope of the study to look at a larger set of countries and found that the evidence was not only limited to Africa. Their results supported hypotheses that group fractionalization has a negative impact on growth, as it leads to rent-seeking behavior and reduces the consensus for public goods. They returned to use the ELF index as a control variable in Easterly and Levine 2004 while studying the relationship between aid, policy and growth, through which it became a standard control variable in research explaining variations in economic performance across countries.

The ELF index utilizes a combination of ethnic and linguistic variables to define ethnolinguistic groups, which means that the specific effect of only linguistic diversity is difficult to estimate. It has therefore been developed further into a pure linguistic fractionalization index by Alesina et al. (2003) using data on shares of languages spoken as native languages in 201 countries. They studied the determinants of the quality of institutions and of growth using variables on linguistic, ethnic, and religious fractionalization. They found that linguistic fractionalization is an important determinant of economic performance and growth, but that its correlation with other explanatory variables makes estimations of the size of its effect—as well as its interpretation—difficult. The results indicated a negative relationship between linguistic fractionalization and economic success.

Another extension of the fractionalization index is Greenberg's (1956) GI index. Desmet, Weber and Ortuño-Ortín (2009) studied how the introduction of distances into measures of diversity affected the significance of the results when using linguistic diversity measures as determinants of redistribution. Contrary to the findings of Easterly and Levine (2004) and Alesina et al. (2003), linguistic fractionalization failed to provide significant results in Desmet, Weber and Ortuño-

Ortín's study of a broad cross-section of countries. Including distances through GI did however yield both statistically and economically significant results.

Researchers focusing on language and civil conflict commonly use a second class of linguistic diversity indices, namely polarization indices, which attempt to explain outcomes in the dependent variable through group size disparities. Marta Reynal-Querol (2002) developed the RQ polarization index to analyze how social cleavages can explain the incidence of ethnic civil war. Montalvo and Reynal-Querol (2005) studied the effect of linguistic fractionalization and polarization on civil conflict for a sample of 138 countries surveyed over the timeframe of 1960–1999, and found that the fractionalization measure possessed weak explanatory power for civil conflict, whereas polarization measures provide both statistically and economically significant results. An increase in polarization was also found to have an indirect negative impact on growth through increases in the prevalence of civil conflict and reductions in the rate of investment.

Similarly, Esteban and Ray (1994) studied the effect of group differences and their effect on social tension and developed a polarization measure, the ER index, which groups individuals into clusters so that the members of each cluster possess similar attributes but are dissimilar to other clusters. The measure also included distances between the groups. They found that polarization was closely related to social tension and unrest.

Desmet, Weber and Ortuño-Ortín (2009) found that the ER and RQ measures provide similar results, as do ELF and GI. ER, which accounts for linguistic distances, provides significant results when estimating its effect on redistribution, whereas RQ does not. They also studied an additional index, which is a combination of two classes of diversity indices. That is the peripheral heterogeneity (PH) index. It is similar to GI but accounts only for distances between a country's central language and its peripheral languages, not distances between the peripheral languages themselves. (Desmet, Weber and Ortuño-Ortín 2005) Desmet, Weber and Ortuño-Ortín (2009) found this index to provide statistically significant results on the same level as GI and ER and argued that it was an advantageous alternative to the more widely used GI when ample lexicostatistical data is available, as it requires fewer computations.

As described above, the majority of the studies on language economics have found a negative relationship between linguistic diversity and economic performance. This can cause countries to wish to standardize language use in order to avoid the negative externalities of linguistic diversity.

This means that a set of official languages are chosen to be used for administrative, educational, and legal purposes to increase efficiency and reduce translation costs and other expenses related to a multilingual language policy. Standardization can however be problematic. Ginsburgh and Weber (2014) (Ginsburgh, Ortuño-Ortín and Weber 2005) present the problem of disenfranchisement, which arises when a country seeks to standardize and reduce the number of official languages. By excluding languages, a country can reduce the costs for maintaining languages but it can also create a sense of alienation among groups that do not speak the official language. They may not have equal access to public information and laws, and such a policy could thereby threaten to limit their involvement in the society's social, economic, and political life. Countries that seek to standardize language use must therefore weigh the benefit of linguistic homogeneity against the social costs of potentially alienating disenfranchised groups. Disenfranchisement is however reduced if a disenfranchised individual speaks the core language as a non-native language or if she speaks a language that is linguistically similar to the core language (Ginsburgh, Ortuño-Ortín and Weber 2005).

There are also studies that dispute the notion that diversity is an impediment to economic success. Florida (2002) compared creativity and innovation rankings with diversity ranks for large, medium, and small-sized cities in the United States and found that there is a strong positive relationship between diversity and innovation, especially for metropolitan areas. Firms, cities and countries failing to adapt to the creative age and adopt new organizational and cultural patterns will lag behind. Florida and Gates (2001) studied determinants of technological success – finding that the metropolitan areas with the highest concentration of foreign-born residents were also the top high tech regions. They obtained highly significant results using diversity to predict high-technology growth.

Further contesting the notion that diversity has an overall negative impact on economic success is Collier (2001). He studied the effect of fractionalization in democracies and dictatorships and found it having only a detrimental effect in the latter case. He also found that the results differed between public and private sectors, implying that public sector performance is impacted negatively by fractionalization whilst the opposite occurs for private sector performance. He also studied polarization measures and found that the implications of dominant groups may be weaker than other researchers claim.

Further support for the positive effects of diversity is given by Parrotta, Pozzoli and Sala (2014), who used employer-employee data for Danish firms between 1995 and 2007 in order to study the impact of workforce diversity on firms' exporting performance. With the hypothesis that firms that manage a diverse workforce are better able to operate in multicultural environments, they found significant support for the notion that more diversified firms perform better on the international market. Exports require knowledge about foreign markets and customer bases, and keeping a diverse workforce may provide employees with the skills needed to operate in multicultural environments, which may in turn facilitate firms' presence on international markets.

## 3.0 Hypotheses

Based on the literature presented above, we derive two hypotheses regarding the question of the effect of linguistic diversity on productivity. First of all, we have expectations relating to the advantage of the linguistic measures themselves. Communicative ability is not in a binary fashion dependent on language. That is, the very fact that individuals speak in what is defined as separate languages does not help to fully predict the communicative prospects between individuals. These languages can be varying in their differences between one another – and a greater deal of linguistic similarity may alleviate the hampering of communication brought upon by the linguistic differentiation. For these reasons we hypothesize that the measure of linguistic diversity that carries the greatest degree of statistical significance is an index that accounts for linguistic distances between languages.

The second hypothesis regards the focus of our study, namely the relationship between linguistic diversity and economic productivity in economically prosperous countries. We expect there to be a positive correlation within the given sample between linguistic diversity and economic performance. This belief is motivated by the composition of the sample, consisting of the developed industrial member-states of the OECD. The theoretical basis for such a hypothesis is that countries with more diverse populations can achieve higher levels of productivity through increased innovation brought upon by the presence of a more diverse range of backgrounds and ideas. There could however be an issue in that it may be the case that this relationship goes two ways so that there is reversed causality, which could possibly interfere with our analysis. This would mean that not only can linguistic diversity affect economic prospects, but that economic outcomes can also affect linguistic diversity through—for instance—higher migratory pull toward productive countries.

# 4.0 Data and Method

## 4.1 Data

To study the effect of linguistic diversity on economic performance, we construct a database of cross-sectional data consisting of economic, linguistic and demographic variables for the 34 member countries of the OECD. Previous research has mainly focused on larger heterogeneous sets of countries from the whole world or on geographically restricted samples. The majority of these studies have found a negative relationship between different economic performance measures and linguistic diversity. But it has not been studied whether or not this relationship varies depending on a country's level of economic success. In order to focus on the effect on economic performance for economically successful countries only, we are restricting our sample of countries based on this parameter. There are a number of different ways to do this using different indicators of economic performance such as, for example, GDP growth and investment levels. These kinds of restrictions, however, require an arbitrary definition of what defines economic success. Different indicators must be given weights in order of importance and a cut-off level must be decided in order to exclude less successful countries.

To avoid these issues, we study the 34 countries which all fulfill the requirements to be part of the OECD. They are some of the world's most successful economies with fairly similar societal sentiments in regards to democracy and economic policy. But in terms of linguistic diversity there are big differences. If linguistic diversity has a negative effect on economic performance one would expect the most successful economies to be linguistically homogenous. This is however not the case. Consider for example the case of the United States and Japan. One is highly linguistically diverse with many large language groups and the other is the opposite, with largely only one major language spoken. Yet they are both two of the strongest economies in the world. Using a sample of OECD countries, we can study the effect of linguistic diversity on economic performance without needing to make any additional assumptions about what defines economic success. Furthermore, the OECD provides detailed and extensive data for all its member countries – including data on capital stocks, labor, and output, which are used to calculate Total Factor Productivity. The database also includes important control variables such as demographic and educational data.

Linguistic variables have been obtained from a number of sources. Shares of speakers of different languages per country is based on data that Alesina (2003) collected from Encyclopedia Britannica (2001). Some adjustments have been made, particularly when it comes to various linguae francae, certain languages known by most and switched to when the speakers do not share the same native tongue. Alesina dropped languages that were linguae francae, which could be misleading as it makes linguistically homogenous countries appear more diverse than they really are. In these cases, a number of sources were considered in order to obtain correct language shares – such as Encyclopedia Britannica, the CIA World Factbook and Ethnologue. Also, in the dataset from Alesina, some countries' languages were expressed in more detail than others, so for the sake of consistency we restricted the data to only include languages spoken by more than .5 % of each country's population. For simplicity, as well as for the reason of limited data, we must also assume that each group only speaks its native language. This is of course very restrictive as many people are either multilingual or otherwise able to communicate with individuals from other linguistic groups. For instance, some groups speak languages that are internally very close—such as Norwegian and Swedish. This is why linguistic distances are important. Data on linguistic distances are obtained from Fearon's (2003) compilation of language trees gathered from Ethnologue. Alesina has provided data on religious fractionalization, which is used as a control variable.

## 4.2 Method

### 4.2.1 Measuring linguistic diversity

There is a large number of models being used in existing literature when measuring linguistic diversity. We will follow Desmet, Weber and Ortuño-Ortín's (2009) index of social effective antagonism and consider the most prevalent linguistic diversity models below.

Social effective antagonism was initially modelled by Esteban and Ray (1994) combining the concepts of identification and alienation. Consider a country with a population $N$ that consists of $K$ different linguistic groups, indexed as $j = 1,2, \dots, K$.

$$N = \sum_{j=1}^{K} N_j$$

(2.0)

Then the share of the population that speaks language $j$ is:

$$s_j = \frac{N_j}{N} \tag{3.0}$$

Where $N_j$ is the population that speaks language $j$.

Identification signifies that an individual belonging to a certain group identifies with other members of the same group. Esteban and Ray (1994) propose that there is a degree of identification dependent on the size of the group. This degree can be denoted by ${s_j}^{\alpha}$ where alpha is either positive—to signify that a larger group contributes to a higher level of identification—or zero to allow for identification to be independent of group size.

Related to the sense of identification to an individual's own group's members is the alienation towards individuals of other groups. By identifying with a member in group $j$, an individual will be unable to identify with a member of group $k$, and may thereby feel a sense of alienation. This sense is increasing with the distance between groups $j$ and $k$, expressed as $\tau_{jk}$ (Desmet, Weber and Ortuño-Ortín 2009). To understand the effect of distances, consider that some pairs of languages represented within the same country are very similar—such as Russian and Belarusian in Belarus—whilst others are radically dissimilar, such as Basque and Castilian Spanish spoken in Spain. Most likely, alienation between language groups should then be greater in Spain than in Belarus – at the very least on purely lexical grounds.

Together, the two forces of identification and alienation create antagonism given by $s_j^{\alpha}\tau_{jk}$ between an individual of group $j$ and an individual of group $k$. Given the shares of people belonging to these two groups, the effective antagonism between the groups is given by $s_k s_j^{1+\alpha}\tau_{jk}$ and the social effective antagonism of the country is the sum for all pairs of groups within that country, referred to as the A-index.

$$A(\alpha,\tau) = \sum_{j=1}^{K}\sum_{k=1}^{K} s_k s_j^{1+\alpha}\tau_{jk} \tag{4.0}$$

Distances $\tau_{jk}$ are computed using language trees. Just like people are genetically related to their ancestors, languages are related to their ancestors through history. To calculate distances, each pair of languages are compared against each other to find at what stage the two languages diverted from their common language tree, and assigned points on the basis of common nodes between language groups $j$ and $k$, expressed as $CN_{jk}$ and ranging from 0 to 15 (Fearon 2003). This can be converted into a continuous measure of linguistic distance on the interval 0 to 1 using the following formula (Spolaore and Wacziarg, 2015).

$$T_{jk} = \sqrt{\frac{(15 - CN_{jk})}{15}} \tag{5.0}$$

There is also a second approach to measuring language distances that we will not use. It is based on lexicostatistical distances and uses the similarities between two languages to define the distance between them, but excludes similarities caused by accident or from borrowing. Instead it focuses on only one dimension—cognate words—which are words that share a common ancestry. A list of 200 basic words, or more precisely '*meanings*', are commonly used since any more would be difficult to manage as well as because the meanings must be existent in all languages of interest. The words for each language pair are compared and lexicostatistical percentages are established using cognates shared by each list of meanings. It is thus a more detailed approach compared to using languages trees since it includes not only linguistic ancestry but also the ancestry of individual words. Tree-based distances are however equally significant and easier to compute when there is a large number of languages to compare. The lexicostatistical approach is based on only one dimension—common roots of words in the vocabulary of languages. The tree-based approach includes this as well, but in addition thereto also syntax, grammar and phonology. It is also available for almost all languages.

Desmet, Weber and Ortuño-Ortín (2009) present three types of distance matrices to be used when measuring linguistic diversity. The first, $T$, is the continuous measure described above. The second, $T_d$, is dichotomous with the linguistic distance $\tau_{jk} = 1$ for all $j \neq k$. This means that measures using $T_d$ are independent of distance so that the alienation experienced by groups does not increase with distance between languages. The third matrix is $T_c$, which is the distance between the central language group, $c$, and the minority languages. It does not incorporate distances between the minority languages themselves.

Similarly to Desmet, Weber and Ortuño-Ortín (2009), we will consider five linguistic diversity measures that attempt to describe the effective social antagonism between language groups of a country. These are the most frequently occurring in existing literature on language and economics. Two of them are fractionalization measures, two are polarization measures and the fifth is a measure that transcends the boundaries of these definitions. They are all expressed as special cases of the A-index below.

a. Linguistic Fractionalization - LF

$$A(0, T^d) = 1 - \sum_{j=1}^{K} s_j^2 \tag{6.0}$$

Fractionalization indices measure the probability of two randomly chosen individuals belonging to different groups. Similar to the well-known ELF index for ethnolinguistic fractionalization (Atlas Narodov Mira 1964, Easterly and Levine 1997, Alesina et al. 1999) but considering linguistic groups rather than ethnolinguistic ones – we obtain the LF measure. It has an $\alpha = 1$ and uses $T_d$ so that distances between groups are not taken into account.

b. Greenberg Index - GI

$$A(0, T) = \sum_{j=1}^{K} \sum_{k=1}^{K} s_k \, s_j \tau_{jk} \tag{7.0}$$

The GI index was introduced by Greenberg in 1956. It includes the element of distances $T$ between languages and can be described as the expected distance between two randomly chosen individuals. It has an $\alpha = 0$.

c. Reynal-Querol - RQ

$$A(1, T^d) = \sum_{j=1}^{K} s_j^2 \, (1 - s_j) \tag{8.0}$$

The RQ index is a polarization index developed by Reynal-Querol (2002). While fractionalization increases when there are many small groups, polarization is maximized when there are few groups of equal size. This measure does not take into account distances between languages and it has an $\alpha = 1$.

d. Esteban and Ray - ER

$$A(1,T) = \sum_{j=1}^{K}\sum_{k=1}^{K} s_k s_j^2 \tau_{jk} \tag{9.0}$$

This is a polarization measure that does control for distances using the continuous distance matrix $T$ and an $\alpha = 1$ (Esteban and Ray, 1994). It is perfectly correlated with the RQ index when distances between all language groups in a country are the same.

e. Peripheral Heterogeneity - PH

$$A(0,T^c) = 2\sum_{j=1}^{K} s_j\, s_c \tau_{cj} \tag{10.0}$$

The Peripheral Heterogeneity index (Desmet, Weber and Ortuño-Ortín, 2005) is similar to GI but controls only for distances between the central (majority) group and the peripheral minority groups, $T_c$, which means that the distances between the minority groups themselves are not taken into account. It has an $\alpha = 0$.

**4.2.2 Measuring productivity**

Total Factor Productivity (TFP) is used in this paper as the main measure of economic performance. The reason for choosing to look at productivity rather than output is that there are a number of conflicting hypotheses as to how other measures of social divergence affect economic performance (Grafton, Knowles and Owen, 2004). One example is the hypothesized effects of ethnic diversity on civil conflict and how this impedes the development of institutions and policies. This type of mechanism is different from the effects in which we are interested, such as disruptions of communication, growing networks, and the pooling and exchange of knowledge and ideas. The difference lies in the fact that the hypothesized effects on institutions and policies have a direct effect on accumulation of production factors such as labor and capital, whereas the communicative effects only affect productivity. To avoid spillovers from conflicting theories we will focus on TFP rather than any output measure in order to capture the effect of linguistic diversity on productivity, as this measure is not affected by the accumulation of production factors.

TFP is a residual that measures the shift in the production function (Hulten 2000). It can be defined as the change in output given a certain level of inputs, or as the part in the output of a country which cannot be explained through its labor and capital levels (Graft, Knowles and Owen, 2004). Its variation is what causes countries with equal resources to achieve unequal production outcomes. TFP is often said to represent innovation in technology, but it also includes changes in attitude and behavior as well as organizational changes and omitted variables (Hulten 2000). These different factors cannot be individually extracted from the bundle that is TFP. It is an oft-criticized measure due to its role as a residual or even—as some term it—"*a measure of our ignorance*" (Hulten 2000) and some theorists claim that it suffers from sins of omission as it fails to measure gains in product quality and costs of growth. However, under the right assumptions TFP is a valid measure of the shift in the production function and it succeeds to outperform all alternative measures. Measurement procedures should be decided upon based on the theory at hand and this implies that we should employ a measure that provides the best fit for our hypothesis. Using TFP, we may not be able to fully extract the effect of linguistic diversity, as it tends to understate the importance of productivity changes for output growth (Hulten 2000), but it may be the closest thing thereto.

Estimates for TFP are obtained using the Cobb-Douglas production function:

$$Y_i = A_i L_i^{\beta} K_i^{\alpha} \tag{11.0}$$

Where $Y$ is output expressed as GDP, $L$ is labor input given by total number of hours worked in a year, $K$ represents capital stocks, and $A$ denotes TFP. $\alpha$ and $\beta$ are the output elasticities of capital and labor given by the available technology. Using data on Y, L and K and assuming the values for the constants $\alpha$ and $\beta$, TFP can be derived. We will herein assume the values $\alpha = \frac{1}{3}$ and $\beta = \frac{2}{3}$. This is a simplification as these values can vary between countries. It is however supported by Hall and Jones (1999) as their estimates using this assumption were very close to the estimates obtained in Hall and Jones (1996) where the exponents were derived separately for each country. As $\alpha$ and $\beta$ sum up to one, the function has a constant return to scale, which also means that doubling each input will exactly double output.

### 4.2.3 Measurement issues

There is a group identification problem common in linguistic studies as there is no fixed rule for how groups should be arranged or when two language systems should be considered to belong to different language groups. Should for example American English and British English be considered the same language? Should Bosnian and Croatian? What we term a *language* is in fact often a group of dialects, which despite being different, are considered to belong together. Group definitions are often tied to political considerations rather than objective linguistic criteria such as mutual intelligibility. For example, following the division of Yugoslavia, the previously termed Serbo-Croatian language was in many official contexts split into four almost identical *languages* (Bosnian, Croatian, Serbian and Montenegrin) to reflect the new political situation. Further, not all researchers use the same criteria to separate languages from dialects, which creates a comparability problem. Alesina et al. (2003) and Fearon (2003) both use the extensive and detailed data from the Ethnologue project, which allows distances between languages to be computed. This detail forces the user to make own assumptions regarding group associations. For example, Alesina et al. (2003) chose only the roughest disaggregation, which separates languages but ignores any dialectal differences. Fearon (2003), however, considers separate any two languages with a distance larger than zero – which allows 291 languages in Mexico compared to the 37 language groups Alesina et al. propose. We follow Alesina's example as this allows for better analysis of communicative effects of linguistic diversity. Diversity measures at lower levels of disaggregation are more significant determinants of productivity and growth, whereas high levels are more suitable for studies on civil conflict (Ginsburg and Weber, 2016). If we were to use higher levels of aggregation such that groups are separated on a dialectal level, the measures that do not include distances would be very misleading. Consider for example a country where there is a large number of different dialects spoken whose distances $\tau_{jk} \neq 0$ but where they are similar enough to be understood by everyone. Then, for measures such as LF and RQ, these dialects would be treated the same as any two different languages making countries appear more linguistically diverse than they should. This is especially problematic if the amount of detail in the database differs between countries.

Another issue when studying data on languages over a large cross-section of countries is that most data tends to be based on national censuses. This means that it may have been compiled differently and might be available for different years in each country. For instance, one country may have data available for 2003 that includes all spoken languages including dialects and another may have data for only 2007 that only includes major linguistic groups. The group identification problem is as mentioned solved by instituting a lower threshold as to what size a group is required to have in order to be included. The time problem is however difficult to evade. Linguistic diversity data is scarcely available in time-series, which means that utilizing cross-country observations is the only real option—even if it essentially implies surveying different time periods for different countries. This, however, should not be taken as cause for any greater concern, as studies (Mauro, 1995) show that this variable displays very little variation over time.

## 5.0 Results

To study the impact of language diversity on economic success, five of the measures of linguistic diversity that are most prevalent in existing literature have been compared. It is of interest to find whether or not they provide similar results and which measure that provides results at the highest level of statistical significance. The measures are the Linguistic Fractionalization (LF), GI, RQ, ER and PH indices described above.

Table 1 displays the result of five simple OLS regressions with the implied total factor productivity of the OECD member countries as the dependent variable and the various measures of linguistic diversity as independent variables. The coefficients are positive for all five measures – indicating a positive relationship between linguistic diversity and total factor productivity. However, only LF and RQ provide results significant at the 1% level. The first one is a fractionalization measure and the other a polarization measure and neither of them account for lexical distances. All other measures provide insignificant results.

**Table 1. Simple regressions**

| VARIABLES | (1)<br>lImpTFP | (2)<br>lImpTFP | (3)<br>lImpTFP | (4)<br>lImpTFP | (5)<br>lImpTFP |
|---|---|---|---|---|---|
| lLing_frac | 0.0644*** | | | | |
| | (0.0215) | | | | |
| lGI | | 0.0336 | | | |
| | | (0.0209) | | | |
| lRQ | | | 0.0680*** | | |
| | | | (0.0235) | | |
| lER | | | | 0.0166 | |
| | | | | (0.0114) | |
| lPH | | | | | 0.0339 |
| | | | | | (0.0215) |
| Constant | -0.0229 | -0.0319 | 0.00446 | -0.0385 | -0.0418 |
| | (0.0226) | (0.0333) | (0.0318) | (0.0321) | (0.0281) |
| Observations | 34 | 34 | 34 | 34 | 34 |
| R-squared | 0.219 | 0.063 | 0.106 | 0.078 | 0.054 |

Standard errors in parantheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

## 5.1 Variables

These results indicate that there may be a positive correlation between productivity and linguistic diversity for industrial countries such as the member states of the OECD. In order to be able to study the effect closer and draw any far-reaching conclusions, it is necessary to include additional control variables.

When predicting the productivity of a nation, diversity of language is not the only diversity factor that is believed to have an influence thereon. Both religious and ethnic fractionalization have often been linked to impacts on productivity (Alesina et al. 2003, Grafton Knowles and Owen 2004). These fractionalization indices—including on language—are usually taken as exogenous given that they remain stable over long time periods, and for comparability with other studies they will necessarily have to be controlled for.

Expanding beyond religious fractionalization, something which is often tied to the prevalent religion(s) of a country or region is culture. Much of Western Europe, for instance, has since the Protestant Reformation been largely Protestant, which may be said to have made distinct the culture of those areas and countries – and indeed there is an abundance of material written on the subject of, for instance, the *Protestant Work Ethic* and its economic implications. While culture is the actual factor that one would seek to include, a general trans-national cultural viewpoint is often indicated by its general religious adherence (e.g. Catholic and Protestant European nations). For this purposes, we include in the regression a number of dummy variables expressing whether *I)* a country is traditionally majority-Protestant, *II)* a country is traditionally majority-Catholic, *III)* a country is traditionally majority-Islamic, or if *IV)* a country is traditionally majority-Jewish. In regards to why these four specific religions—Protestantism, Catholicism, Islam, and Judaism—are included, the reasoning is that all the sampled countries have historical majorities of either one of these four particular faiths or Shinto in the case of Japan, making the inclusion of other religions herein superfluous. The dropping of the final religion—Shinto—as a dummy variable is done to have one less dummy variable than possible categories. This is also consistent with the variables used by La Porta et al. (1999) and Desmet, Weber and Ortuño-Ortín (2009).

Generally, education is often included as a factor whose effect on productivity is positive. While education and educational standards may differ wildly between nations, the average years of schooling undertaken by a country's residents typically serves as a fairly good indicator of the overall educational level. In less developed countries, specific educationally obtained skills such as literacy may be utilized but in industrial countries—such as the OECD—there is generally near-total literacy rates as it is commonly achieved at a very basic educational level. This means that a specific skill-measure such as literacy would do little to differentiate the educational levels of two industrial nations. For this reason, a more nuanced picture that could perhaps be obtained through comparing specific skills in greater detail is very difficult to achieve, causing the average years of education to serve as our indicator of educational attainment in a country. Indeed, in the regressions run, the independent variable of years of education is statistically significant, lending credence to its relevance as such an appraiser of national scholastic development.

Other control variables included in our regressions are population as well as dummy variables for the general origin of countries' legal systems (e.g. countries with Common Law systems having an English legal origin).

We thus formulate the regression as follows, where *Language_diversity* is a generic representation of either one of the five such indices that we are investigating.

$$\begin{aligned} lImpTFP_i = \beta_0 &+ \beta_1 lLanguage\_diversity_i + \beta_2 lRel\_frac_i + \beta_3 Yrseduca_i + \beta_4\_POP_i \\ &+ \beta_5 Protestant_i + \beta_6 Catholic_i + \beta_7 Judaism_i + \beta_8 Islam_i + \beta_9 English_i \\ &+ \beta_{10} French_i + \beta_{11} Scandinavian_i + \beta_{12} Socialist_i + \varepsilon_i \end{aligned} \tag{12.0}$$

The logarithmic functional form allows for percentage-based comparisons, e.g. the logarithmic forms enabling translations such as a percentage-change in the dependent variable corresponding to a percentage change in an independent variable determined by the value of its coefficient estimator.

Breusch-Pagan tests for heteroscedasticity for each of the five regressions (the 11.0 equation takes on five different forms through the five different language diversity indices) are run in Stata through the *hettest* command.

**Table 2. Heteroscedasticity test**
**$H_0$:** Constant variance of error terms, i.e. homoscedasticity. $E(u^2|x_1, x_2) = E(u^2) = \sigma^2$

| Measure of $Language\_diversity$ | ${\chi^2}_{12}$ | $Prob > \chi^2$ |
|---|---|---|
| *Linguistic Fractionalization Index* | 5.77 | .9273 |
| *Greenberg Index* | 6.04 | .9140 |
| *Reynal-Querol Index* | 5.87 | .9225 |
| *Esteban-Ray Index* | 7.99 | .7857 |
| *Peripheral Heterogeneity Index* | 6.07 | .9126 |

Based on this we cannot reject the null-hypothesis of homoscedasticity, and therefore we do not require heteroscedasticity-robust standard errors.

## 5.2 The regressions

The regressions are run separately for each of the five measures of language diversity.

### 5.2.1 Linguistic Fractionalization Index regression

Table 3 displays the results of our regressions of TFP on the Linguistic Fractionalization index (LF). The first column is similar to what was shown in Table 1 and includes the LF index as the only independent variable. It has a positive coefficient and exhibits statistical significance at the 1 % level. In Column 2 the *Years of education* and in Column 3 the *Religious fractionalization* and *Population* variables are added. In both of these regressions, all variables other than *Population* are highly significant at the 1 % level. The coefficient for LF decreases as the new independent variables are added but remains positive. In Column 4 we add a fourth independent variable—*Ethnic fractionalization*. It does not provide a better fit in terms of R-squared and removes the significance from our variable of interest, LF. This is likely due to the high internal correlation between linguistic and ethnic fractionalization, being 0.8382. The reason for this is that ethnic groups are often, but not always, differentiated in terms of language. If we were to look at the United States, for instance, there are Hispanic and Chinese groups that speak Spanish and

Chinese, respectively, but there are also ethnic groups that share a common language between them—such as most non-Hispanic Whites and African-Americans. As the *Ethnic fractionalization* variable is insignificant, highly correlated with LF, and does not add anything new to our model it is kept out in the following regressions.

Column 5 adds dummy variables for the legal origin of different countries. These are *English legal origin*, *French legal origin*, *Scandinavian legal origin* and *Socialist legal origin*. The last category *German legal origin* is kept out of the regression. The inclusion of these dummies reduces the coefficient for LF slightly and its significance level decreases to the 5 % level. It improves the fit of the model, however, which is consistent with theory as differences in legal origins cause differences in institutions and property rights and thereby the prospect of economic success (Levine 2005, Olson 1996). For Column 6, previously described cultural-religious dummy variables are added to the regression. For comparability, this is also the regression that is closest in form to earlier works such as La Porta et al. (1999), Alesina et al. (2003) and Desmet, Weber and Ortuño-Ortín (2009). We find that adding these variables makes LF insignificant at any interesting level. However, if we remove the legal origin dummies as in Column 7, LF is again statistically significant but at a lower level than in Column 5—being 10 %. The reason for the lower significance in Column 6 may be multicollinearity. There is for example strong correlation between the *Protestant* dummy and the *Scandinavian Legal Origin* dummy. A potential problem with such a collinearity issue is that, although the explanatory potential of the regression model as a whole may remain satisfactory, the estimators of the independent variables' coefficients are each on their own unreliable. Therefore, any future analysis will be based on the results displayed in Column 7, which is the model whose fit is also the best.

Dropping these above-mentioned variables produces the following regression model:

$$lImpTFP_i = \beta_0 + \beta_1 lLanguage\_diversity_i + \beta_2 lRel\_frac_i + \beta_3 Yrseduca_i + \beta_4\_POP_i + \beta_5 Protestant_i + \beta_6 Catholic_i + \beta_7 Judaism_i + \beta_8 Islam_i + \varepsilon_i \tag{13.0}$$

Further—as shown in the regression results in Table 3—we learn that the impact of LF on TFP is robust for all regressions above, except for when we have strong multicollinearity between the independent variables. It is significant at a 1 % level when religious affiliation and legal origin is not taken into account and significant at a 5 and 10 % level, respectively, when the two variables are included. The LF coefficient is positive for all regressions but decreases as more variables are added, ranging from 0.0644 to 0.0321. For the elected model in the rightmost column, we find that a one-percent increase in a country's linguistic fractionalization corresponds to a 3.21 % rise in TFP.

**Table 3. Full Regression Results LF and TFP**

| VARIABLES | (1) lImpTFP | (2) lImpTFP | (3) lImpTFP | (4) lImpTFP | (5) lImpTFP | (6) lImpTFP | (7) lImpTFP |
|---|---|---|---|---|---|---|---|
| lLing_frac | 0.0644*** | 0.0588*** | 0.0517*** | 0.0494 | 0.0428** | 0.0241 | 0.0321* |
| | (0.0215) | (0.0201) | (0.0169) | (0.0323) | (0.0161) | (0.0177) | (0.0173) |
| lRel_frac | | | -0.115*** | -0.115*** | -0.118*** | -0.113** | -0.127*** |
| | | | (0.0295) | (0.0300) | (0.0295) | (0.0516) | (0.0373) |
| Yrseduca | | 0.0198** | 0.0431*** | 0.0432*** | 0.0434*** | 0.0476*** | 0.0399*** |
| | | (0.00784) | (0.00889) | (0.00916) | (0.0106) | (0.0110) | (0.00851) |
| _POP | | | 1.23e-10 | 1.22e-10 | 6.98e-11 | 6.23e-11 | 1.55e-10 |
| | | | (1.54e-10) | (1.58e-10) | (1.49e-10) | (1.48e-10) | (1.44e-10) |
| Protestant | | | | | | 0.0765 | 0.0992*** |
| | | | | | | (0.0532) | (0.0295) |
| Catholic | | | | | | 0.0726** | 0.0702** |
| | | | | | | (0.0291) | (0.0290) |
| Judaism | | | | | | 0.0477 | 0.0837 |
| | | | | | | (0.0675) | (0.0567) |
| Islam | | | | | | 0.0952 | 0.0185 |
| | | | | | | (0.129) | (0.0854) |
| English | | | | | 0.0608** | 0.0441 | |
| | | | | | (0.0250) | (0.0427) | |
| French | | | | | 0.0376 | 0.0358 | |
| | | | | | (0.0245) | (0.0331) | |
| Scandinavian | | | | | 0.0397 | 0.0192 | |
| | | | | | (0.0263) | (0.0645) | |
| Socialist | | | | | -0.0398 | -0.0616 | |
| | | | | | (0.0503) | (0.0490) | |
| lEthnic_frac | | | | 0.00283 | | | |
| | | | | (0.0338) | | | |
| Constant | -0.0229 | -0.251** | -0.580*** | -0.581*** | -0.620*** | -0.740*** | -0.641*** |
| | (0.0227) | (0.0926) | (0.115) | (0.117) | (0.138) | (0.144) | (0.114) |
| Observations | 34 | 34 | 34 | 34 | 34 | 34 | 34 |
| Adj. R-squared | 0.194 | 0.310 | 0.517 | 0.500 | 0.593 | 0.620 | 0.622 |

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

### 5.2.2 Greenberg Index regression

Running the same regressions with GI as our independent variable of interest gives us the results displayed in Table 4. This shows how the effect of fractionalization on TFP differs when distances are taken into account. For instance, we see that significance is only strong in two out of seven cases: in Column 2 and 5. The first one is when we add *Years of education* into the regression. This makes all independent variables in the regression significant and specifically it gives GI a statistical significance level of 10 %. However Columns 3 and 4 both provide insignificant results. When we add the legal origin dummies GI is once again significant at the 10 % level. Overall we can see that the effect of GI on TFP is not robust and that it fails to be significant even at the 10 % level for most of the regressions. This is the opposite of what we might expect and the reverse of what Desmet, Weber and Ortuño-Ortín (2009) found in a very similar study. They found that GI was able to find significant results where LF could not and that both measures showed negative coefficients. The difference is surprising but not implausible. Their study used data for over 200 countries where we look only at OECD countries, it used different definitions and computations for groups and distances and focused on redistribution rather than TFP. One explanation for why distances are of less importance here is that for our set of countries educational levels are high which facilitates communication through non-native languages, or world languages such as English, for people belonging to different language groups.

**Table 4. Full Regression Results GI and TFP**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| VARIABLES | lImpTFP | lImpTFP | lImpTFP | lImpTFP | lImpTFP | lImpTFP | lImpTFP |
| lGI | 0.0336 | 0.0330* | 0.0274 | 0.000839 | 0.0287* | 0.0161 | 0.0243 |
| | (0.0210) | (0.0193) | (0.0165) | (0.0217) | (0.0158) | (0.0158) | (0.0150) |
| lRel_frac | | | -0.120*** | -0.117*** | -0.120*** | -0.111** | -0.129*** |
| | | | (0.0323) | (0.0312) | (0.0317) | (0.0525) | (0.0379) |
| Yrseduca | | 0.0222** | 0.0462*** | 0.0462*** | 0.0445*** | 0.0484*** | 0.0404*** |
| | | (0.00842) | (0.00967) | (0.00932) | (0.0114) | (0.0113) | (0.00862) |
| _POP | | | 9.38e-11 | 9.36e-11 | 0 | 0 | 1.34e-10 |
| | | | (1.70e-10) | (1.63e-10) | (1.58e-10) | (1.50e-10) | (1.47e-10) |
| Protestant | | | | | | 0.0875 | 0.117*** |
| | | | | | | (0.0533) | (0.0276) |
| Catholic | | | | | | 0.0818*** | 0.0809*** |
| | | | | | | (0.0278) | (0.0275) |
| Judaism | | | | | | 0.0591 | 0.0984* |
| | | | | | | (0.0675) | (0.0554) |
| Islam | | | | | | 0.111 | 0.0271 |
| | | | | | | (0.131) | (0.0862) |
| English | | | | | 0.0725** | 0.0463 | |
| | | | | | (0.0262) | (0.0434) | |
| French | | | | | 0.0422 | 0.0377 | |
| | | | | | (0.0260) | (0.0339) | |
| Scandinavian | | | | | 0.0529* | 0.0234 | |
| | | | | | (0.0291) | (0.0655) | |
| Socialist | | | | | -0.0352 | -0.0620 | |
| | | | | | (0.0536) | (0.0500) | |
| lEthnic_frac | | | | 0.0459* | | | |
| | | | | (0.0256) | | | |
| Constant | -0.0319 | -0.282*** | -0.624*** | -0.625*** | -0.635*** | -0.756*** | -0.651*** |
| | (0.0333) | (0.0998) | (0.125) | (0.121) | (0.150) | (0.148) | (0.115) |
| Observations | 34 | 34 | 34 | 34 | 34 | 34 | 34 |
| Adj. R-squared | 0.045 | 0.195 | 0.416 | 0.458 | 0.539 | 0.624 | 0.611 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

### 5.2.3 Reynal-Querol Index regression

Table 5 lists the full results when running the same regressions as above using RQ as our linguistic diversity independent variable. RQ is the first of our two polarization measures and does not consider distances. Interestingly it displays very similar results as do LF. It is significant at the 1 % level in the first three regressions, at the 5 % level in Column 5 and at the 10 % level in Column 7. It is not significant when ethnic diversity is included or when religious and legal origin dummies are both included. In terms of fit and coefficient it is also very similar and has an Adjusted R-squared value equal to 0.618 (compared to 0.622) and a coefficient equal to 0.0334 (compared to 0.0321), as shown in column 7—displaying the Reynal-Querol regression results for the previously listed model in equation 13.0.

**Table 5. Full Regression Results RQ and TFP**

| VARIABLES | (1)<br>lImpTFP | (2)<br>lImpTFP | (3)<br>lImpTFP | (4)<br>lImpTFP | (5)<br>lImpTFP | (6)<br>lImpTFP | (7)<br>lImpTFP |
|---|---|---|---|---|---|---|---|
| lRQ | 0.0680*** | 0.0631*** | 0.0554*** | 0.0524 | 0.0464** | 0.0254 | 0.0334* |
| | (0.0235) | (0.0217) | (0.0183) | (0.0355) | (0.0173) | (0.0193) | (0.0188) |
| lRel_frac | | | -0.115*** | -0.115*** | -0.119*** | -0.114** | -0.128*** |
| | | | (0.0296) | (0.0301) | (0.0295) | (0.0518) | (0.0375) |
| Yrseduca | | 0.0204** | 0.0436*** | 0.0437*** | 0.0438*** | 0.0479*** | 0.0403*** |
| | | (0.00784) | (0.00888) | (0.00912) | (0.0106) | (0.0110) | (0.00854) |
| _POP | | | 1.20e-10 | 1.18e-10 | 6.40e-11 | 6.00e-11 | 1.54e-10 |
| | | | (1.55e-10) | (1.58e-10) | (1.48e-10) | (1.48e-10) | (1.45e-10) |
| Protestant | | | | | | 0.0766 | 0.0993*** |
| | | | | | | (0.0533) | (0.0298) |
| Catholic | | | | | | 0.0725** | 0.0705** |
| | | | | | | (0.0294) | (0.0293) |
| Judaism | | | | | | 0.0496 | 0.0873 |
| | | | | | | (0.0675) | (0.0566) |
| Islam | | | | | | 0.0940 | 0.0184 |
| | | | | | | (0.130) | (0.0863) |
| English | | | | | 0.0618** | 0.0445 | |
| | | | | | (0.0249) | (0.0428) | |
| French | | | | | 0.0379 | 0.0359 | |
| | | | | | (0.0244) | (0.0333) | |
| Scandinavian | | | | | 0.0388 | 0.0185 | |
| | | | | | (0.0263) | (0.0647) | |
| Socialist | | | | | -0.0413 | -0.0623 | |
| | | | | | (0.0502) | (0.0491) | |
| lEthnic_frac | | | | 0.00348 | | | |
| | | | | (0.0343) | | | |
| Constant | 0.00446 | -0.231** | -0.563*** | -0.566*** | -0.605*** | -0.733*** | -0.632*** |
| | (0.0319) | (0.0951) | (0.117) | (0.121) | (0.139) | (0.147) | (0.117) |
| Observations | 34 | 34 | 34 | 34 | 34 | 34 | 34 |
| Adj. R-squared | 0.182 | 0.307 | 0.514 | 0.497 | 0.595 | 0.618 | 0.618 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

### 5.2.4 Esteban-Ray Index regression

If we want to account for distances between languages using a polarization measure, we turn to the ER index, whose regression results in their entirety are presented in Table 6 below, with the results for the elected regression model (see equation 13.0) are displayed in its rightmost column. These results differ from what we have seen earlier. ER is not significant even at the 10 % level in the simple regression or in the regression that includes ethnic fractionalization. However, it is strongly significant in all the remaining ones including Column 6. ER is the only measure that provides statistical significance in the regression that includes all independent variables. And for the regression of our choice displayed in Column 7, it is the only variable that is significant at the 5 % level. It does also provide a better fit than the previous regressions with an Adjusted R-squared of 0.653. The coefficient is however lower than before at 0.0190 but is more constant compared to the others, ranging between 0.0107 and 0.0226. In this case the hypothesis posed by Desmet, Weber and Ortuño-Ortín (2009) appears to hold as ER that accounts for distances performs better than RQ, which does not when it comes to explaining the variation in TFP.

**Table 6. Full Regression Results ER and TFP**

| VARIABLES | (1)<br>lImpTFP | (2)<br>lImpTFP | (3)<br>lImpTFP | (4)<br>lImpTFP | (5)<br>lImpTFP | (6)<br>lImpTFP | (7)<br>lImpTFP |
|---|---|---|---|---|---|---|---|
| lER | 0.0166 | 0.0226** | 0.0200** | 0.0107 | 0.0196** | 0.0154* | 0.0190** |
| | (0.0115) | (0.0104) | (0.00875) | (0.0108) | (0.00881) | (0.00823) | (0.00778) |
| lRel_frac | | | -0.119*** | -0.116*** | -0.119*** | -0.114** | -0.134*** |
| | | | (0.0311) | (0.0306) | (0.0307) | (0.0497) | (0.0359) |
| Yrseduca | | 0.0258*** | 0.0492*** | 0.0476*** | 0.0502*** | 0.0515*** | 0.0437*** |
| | | (0.00835) | (0.00933) | (0.00924) | (0.0108) | (0.0102) | (0.00810) |
| _POP | | | 1.04e-10 | 9.60e-11 | 7.92e-11 | 7.45e-11 | 1.47e-10 |
| | | | (1.63e-10) | (1.60e-10) | (1.55e-10) | (1.43e-10) | (1.38e-10) |
| Protestant | | | | | | 0.0943* | 0.115*** |
| | | | | | | (0.0507) | (0.0261) |
| Catholic | | | | | | 0.0802*** | 0.0776*** |
| | | | | | | (0.0258) | (0.0257) |
| Judaism | | | | | | 0.0587 | 0.0816 |
| | | | | | | (0.0629) | (0.0531) |
| Islam | | | | | | 0.0957 | 0.00878 |
| | | | | | | (0.122) | (0.0816) |
| English | | | | | 0.0565** | 0.0289 | |
| | | | | | (0.0265) | (0.0423) | |
| French | | | | | 0.0448* | 0.0348 | |
| | | | | | (0.0250) | (0.0317) | |
| Scandinavian | | | | | 0.0560* | 0.0188 | |
| | | | | | (0.0283) | (0.0622) | |
| Socialist | | | | | -0.0507 | -0.0722 | |
| | | | | | (0.0525) | (0.0472) | |
| lEthnic_frac | | | | 0.0328 | | | |
| | | | | (0.0230) | | | |
| Constant | -0.0385 | -0.313*** | -0.647*** | -0.624*** | -0.690*** | -0.772*** | -0.674*** |
| | (0.0322) | (0.0933) | (0.117) | (0.116) | (0.137) | (0.128) | (0.104) |
| Observations | 34 | 34 | 34 | 34 | 34 | 34 | 34 |
| Adj. R-squared | 0.032 | 0.237 | 0.458 | 0.477 | 0.670 | 0.646 | 0.653 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

### 5.2.5 Peripheral Heterogeneity Index regression

Finally we look at the PH index which is a combination of a fractionalization and polarization index. The full results are displayed in Table 7 below, while the results for the chosen regression model (see equation 13.0) specifically are displayed in its rightmost column Here only one out of seven regressions provide significant results (see Column 5) and even then only at the 10 % level. It appears to be the worst performing measure with GI as a close second. Equaling the results of GI, the fit of the PH regression of our choice (Column 7) shares alongside GI the dubious honor of having the lowest fit, an Adjusted R-squared amounting to 0.610.

**Table 7. – Full Regression Results PH and TF**

| VARIABLES | (1) lImpTFP | (2) lImpTFP | (3) lImpTFP | (4) lImpTFP | (5) lImpTFP | (6) lImpTFP | (7) lImpTFP |
|---|---|---|---|---|---|---|---|
| lPH | 0.0339 | 0.0334 | 0.0280 | 0.000873 | 0.0291* | 0.0100 | 0.0249 |
| | (0.0216) | (0.0198) | (0.0169) | (0.0222) | (0.0162) | (0.0218) | (0.0154) |
| lRel_frac | | | -0.120*** | -0.117*** | -0.120*** | -0.106* | -0.129*** |
| | | | (0.0323) | (0.0312) | (0.0317) | (0.0551) | (0.0380) |
| Yrseduca | | 0.0222** | 0.0463*** | 0.0462*** | 0.0448*** | 0.0491*** | 0.0406*** |
| | | (0.00843) | (0.00967) | (0.00931) | (0.0114) | (0.0115) | (0.00861) |
| _POP | | | 9.40e-11 | 9.36e-11 | 0 | 0 | 1.35e-10 |
| | | | (1.70e-10) | (1.63e-10) | (1.58e-10) | (1.53e-10) | (1.47e-10) |
| Protestant | | | | | | 0.0799 | 0.117*** |
| | | | | | | (0.0570) | (0.0276) |
| Catholic | | | | | | 0.0774** | 0.0812*** |
| | | | | | | (0.0301) | (0.0275) |
| Judaism | | | | | | 0.0561 | 0.0980* |
| | | | | | | (0.0692) | (0.0556) |
| Islam | | | | | | 0.118 | 0.0260 |
| | | | | | | (0.135) | (0.0864) |
| English | | | | | 0.0721** | 0.0449 | |
| | | | | | (0.0263) | (0.0444) | |
| French | | | | | 0.0428 | 0.0391 | |
| | | | | | (0.0260) | (0.0346) | |
| Scandinavian | | | | | 0.0528* | 0.0286 | |
| | | | | | (0.0291) | (0.0678) | |
| Socialist | | | | | -0.0352 | -0.0677 | |
| | | | | | (0.0537) | (0.0525) | |
| lEthnic_frac | | | | 0.0459* | | 0.0115 | |
| | | | | (0.0255) | | (0.0259) | |
| Constant | -0.0418 | -0.292*** | -0.633*** | -0.625*** | -0.647*** | -0.759*** | -0.660*** |
| | (0.0282) | (0.0985) | (0.124) | (0.120) | (0.148) | (0.148) | (0.114) |
| Observations | 34 | 34 | 34 | 34 | 34 | 34 | 34 |
| Adj. R-squared | 0.043 | 0.193 | 0.416 | 0.458 | 0.538 | 0.610 | 0.610 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From the results above we can observe two things. Firstly, distances may not be as important as one may think. The two measures that do not include linguistic distances perform better than two out of three measures that do include them. Secondly, there seems to be some relationship between the two non-distance measures as they exhibit the same significance levels as independent variables are added to the regressions, as well as maintaining similar R-squared values. The same thing goes for the two worst performing measures GI and PH. They both include distances and develop very similarly to each other. This follows Desmet, Weber and Ortuño-Ortín (2009) finding that the performance of diversity measures are dependent on whether or not distances are included so that non-distance measures perform similarly to other non-distance measures and vice versa. However the ER index does not conform to that idea. It does best at explaining the variance in TFP while the other two diversity measures that include distances are the worst. This makes it difficult to come to any clear conclusion regarding the hypothesis that measures accounting for linguistic distances more aptly explain variations in countries' productivity levels.

The higher significance of the ER index and its higher R-squared value motivates the choice thereof as the main measure of linguistic diversity when we estimate its effect on Total Factor Productivity from this point forward. The regression is thus formulated as follows:

$$lImpTFP_i = \beta_0 + \beta_1 lER_i + \beta_2 lRel\_frac_i + \beta_3 Yrseduca_i + \beta_4\_POP_i + \beta_5 Protestant_i + \beta_6 Catholic_i + \beta_7 Judaism_i + \beta_8 Islam_i + \varepsilon_i \tag{14.0}$$

The obtained estimators in Table 6, Column 7 provide the following estimated model to predict TFP:

$$lImpTFP_i = -0.674 + 0.0190 lER_i - 0.134 lRelfrac_i + 0.0437 Yrseduca_i + 1.47 * 10^{-10}\_POP_i + 0.115 Protestant_i + 0.0776 Catholic_i + 0.816 Judaism_i + 0.00878 Islam_i + \varepsilon_i \tag{15.0}$$

And ultimately, what we find—as demonstrated above—is a positive effect with $\hat{\beta}_1 = 0.0190$, the interpretation of which is that a 1 % rise in ER corresponds to a 1.90 % increase in TFP.

It might seem counter-intuitive that the ER measure—a polarization measure—does better at explaining a *positive* effect of linguistic diversity on productivity than any fractionalization measure. Had the language multiplicity effect on TFP been negative, the explanation had followed more intuitively as one could imagine that increased polarization and social antagonism between groups would hamper the productivity of a society. As demonstrated above, the actual effect here is a positive one – and a possible explanation for this could be that—in developed countries—perhaps the case is that polarization causes distinct groups to, to a greater degree, cooperate internally to compete against other groups, an effect that is perhaps stronger when the sense of inter-community polarization is greater – as theorized by Mappes and Puurtinen (2009). It is possible that this could serve as a driver for growth and productivity, thus causing the above observed effect.

# 6.0 Discussion

## 6.1 Correlation between linguistic diversity and productivity

As described above, we find a positive correlation between linguistic diversity—expressed through the ER polarization measure—and Total Factor Productivity. We obtain a statistically significant result with an estimate of the coefficient of 0.0190 for our variable of interest. The interpretation is that a 1 % increase in the ER index corresponds to a 1.90 % increase in TFP. The difficulty in quantifying this relationship may make assumptions about the economic significance of the result ambiguous. Small coefficients in the above regressions may have larger implications when you take a step further, from TFP to output. Even small variances in TFP can ultimately have large impacts on economic reality.

The positive relationship between ER and TFP implies, but cannot prove, a positive effect of linguistic diversity on productivity. See the sub-section below for a discussion on causality. This positive relationship implies that higher levels of linguistic polarization should be positive for economic outcomes. Polarization as expressed by the ER index is maximized when there are only two linguistic groups of equal size and the distance between them is 1. That is, they belong to entirely different language trees for example the Indo-European Italian and Arabic that belongs to the Afro-Asiatic language family. Interestingly, ER fared better than the fractionalization measures in terms of statistical significance and fit. This is important because fractionalization measures are maximized when each individual belongs to their own separate group, which in terms of social effective antagonism would signify high levels of alienation and non-existent identification. Each individual would be alone with their language trait. For the ER index however, both identification and alienation are present as polarization is maximized. This means that the measure that values both mechanisms can explain more of the variance in TFP in this sample than the measure that focuses on alienation. In addition to this effect, it is possible—as argued in the preceding section—that the social mechanisms of inter-community competition could be driving positive productivity effects, which would then indicate a general advantage of polarization indices in highly developed countries, beyond the ER-specific benefits argued here.

On a general language diversity-focused level, there are a number of reasons as to why we find a positive correlation between linguistic diversity—in general terms—and TFP. Overall it implies that the benefits of diversity may outweigh the costs related to the potential impediments to coordination and trust. These benefit may include—but are not limited to—creativity boosts, increased idea sharing, innovation and improved export market adaptation. This phenomenon is related to the focus of our study and our sample choice, namely economically successful countries. We have only been looking at member states of the OECD, which are all high-performing countries that share relatively similar views on democracy and economic policy but differ in terms of linguistic diversity. This allows us to draw certain conclusions about the empirical results of the study.

As we look at this specific group of countries we can relate our results to Florida and Gates (2001) who found a positive relationship between diversity and growth in metropolitan areas through the theorized effects of diversity on innovation. This is important because 30 out of our 34 observations rank in the top 50 in the Bloomberg innovation index. (Bloomberg Innovation Index 2015) We can also find that out of the top ten most innovative countries, nine of these are included in our sample of OECD countries. This tells us that both productivity and diversity should be higher for these countries. Also, if we consider Ashraf and Galor (2013), the positive coefficient tells us that for this sample – the force related to innovation and the extension of the production possibility frontier is stronger than the force related to miscommunication and coordination issues.

Secondly, we can relate the results to the sample countries' positions as export countries. Out of the 34 countries of the OECD, 28 are among the top 50 largest exporters in terms of total (in USD) amount of merchandise exports, and eight of them are among the top ten exporting nations. (CIA World Factbook) This is important as Parrotta, Pozzoli and Sala (2014) found that exporting firms with a diversified workforce perform better in foreign markets and multicultural environments, which agrees well with our findings. This can then help explain the positive relationship between linguistic diversity and productivity that we identified.

## 6.2 Causality issues

As demonstrated above, we observe a generally positive correlation between the level of linguistic diversity within a country and its level of estimated total factor productivity. We cannot determine, however, whether there is a causative relationship, and if so – the direction thereof.0

For example, the relationship we are trying to explain above is that of the effect of linguistic diversity on TFP, not the reverse. Such a scenario could potentially be explained by that higher linguistic fractionalization brings with it a diverse range of cultural perspectives included in language. It is possible that an increase in the diversity thereof could foster a more innovative society, and thus positively affect productivity, through the inclusion of a wider range of perspectives. There exists plentiful research on diversity and its economic implications, and in addition one could imagine that the case could also be that the diversity in ways of thinking could have similar effects. Indeed, the idea that language shapes the way individuals think has been raised by numerous scholars (Boroditsky 2011), and from there one may consider the potential effect thereby of language on productivity as expressed here by total factor productivity.

While the above reasoning assumes a causal relationship where linguistic fractionalization increases productivity levels, the reversed causality is equally plausible. Research on the effect of linguistic diversity on economic outcomes has not taken into deeper consideration that these outcomes may have reversed implications for the development of languages communities. Instead it is sometimes the case that the levels of linguistic diversity are taken as something that is near-given, and independent of economic development (Ginsburgh and Weber, 2016).

This type of effect can be called a feedback mechanism where linguistic factors affect economic decisions, which in their turn affect language dynamics. Consider for example the results presented above: an increase in linguistic diversity corresponds to an increase in economic productivity. Higher productivity is related to higher output, which in turn may either reduce or increase linguistic diversity. In choosing where to migrate, economic opportunities can often be a deciding factor. This means that high-output regions are more likely to attract immigrants, thereby increasing fractionalization. However there is a second effect of economic success on language as well. For a high-output region with a dominant language, speakers of that language are more likely to have power over economic decisions, further increasing the dominance and value of that specific language. High-value languages are more likely to attract more speakers and learners, making them more valuable and decreasing linguistic diversity (Ginsburgh and Weber, 2016). Which of these two effects that is greater is an entirely different question, which we will not attempt to answer here. However we can conclude that there appears to be a spiral effect wherein linguistic diversity affects economic performance which then affects linguistic diversity, and so on.

Although causal relationships are impossible to prove through statistical analysis, we would ideally desire to investigate whether any particular credence could be lent to either of the above mentioned potential causal directions. One way of doing is through Difference-in-Differences.

Initially, in the Difference-in-Differences method we use the same assumptions as in the Ordinary Least Squares model, with the addition of an assumption of parallel trends. Hence, the assumptions are rendered as follows:

1. *Correct specification*
2. *Strict exogeneity* $E(\varepsilon|X) = 0$
3. *Linear independence*
4. *Homoscedasticity* $E(\varepsilon_i{}^2|X) = \sigma^2$
5. *No autocorrelation* $E(\varepsilon_i\varepsilon_j|X) = 0 \; for \; i \neq j$
6. *Normality of errors* $\varepsilon|X \sim N(0, \sigma^2)$
7. *Parallel trends* $\lambda_2 - \lambda_1 \; same \; for \; s = 1, \; s = 2$

The implication of the seventh assumption—concerning parallel trends—is here that the control group as well as the treatment group would be expected to have the same change in total factor productivity over these years in the absence of the increase in language fractionalization. In this case, were one to seek to use Difference-in-Differences on—for instance—a treatment group consisting of a certain nation whose level of linguistic fractionalization underwent a large change, whether in an upward or downward direction, between the two dates observed and a control group of a country whose language fractionalization remained largely unchanged between these points in time, this vital seventh assumption may be argued to be satisfied if the countries' economies are similar in industry composition and otherwise operate in similar macroeconomic climates. If this is not the case, it could be that one nation depends to a relatively large degree on an industry that was disproportionately affected—for instance through large technological changes specific thereto—during the investigated timespan, whereas the other country does not, in which case the two countries would very plausibly have differing productivity developments even in the absence of the treatment.

The problematic aspect of using Difference-in-Differences for the purposes of this causative conundrum is rooted in the very reason language composition tends to change. Within our sample are present the Republic of Korea and Israel, two industrial nations with significant technological sectors and fairly resembling economies. Further, both countries are similar in their respective general market approaches, with the 1980s seeing Israel largely abandoning the state-dominated and planned Socialist economy that had dominated its history since independence in

1948, and introducing market-oriented reforms (Mealem and Melnick, 2009) – aligning with the US-sphere-influenced market-oriented South Korean economy.

**Table 8. Exports 1990**

| *Country* | *High-technology exports 1990 (current US$)* | *As share of GDP* |
|---|---|---|
| Israel | 1,111,523,814 | 2.12% |
| Rep. Korea | 10,936,004,210 | 3.84% |

Data source: World Bank, tables *High-technology exports (current US$), GDP at market prices (current US$)*

With the assumptions fulfilled, we seek to run Difference-in-Differences thereon. The Republic of Korea has a very static level of linguistic fractionalization as it remains very homogenous over time, and contains no relatively significant minority groups (CIA, 2016). Israel, however, saw an abnormally large influx of immigrants following the collapse of the Soviet Union in the early 1990s (Israeli Ministry of Education, 2007), which brought in a very significant Russian-speaking linguistic minority, thus affecting Israel's linguistic fractionalization level. We select a pre-treatment date of 1985, and allow for an end date of 1995. Within this timeframe, we see the Republic of Korea not have its linguistic fractionalization particularly affected, whereas that of Israel rises notably. We therefore, in a bid to investigate a potential causative link between linguistic fractionalization and productivity, designate the rise in linguistic fractionalization the *treatment* – which Israel underwent during this time period.

**Table 9. DiD Data**

| *Country* | *Year* | *Implied TFP* | *Treatment Status* |
|---|---|---|---|
| Israel | 1985 | 2.46 | 0 |
| Rep. Korea | 1985 | 1.00 *[Rep. Korea_1985 set to 1.00]* | 0 |
| Israel | 1995 | 2.92 | 1 |
| Rep. Korea | 1995 | 1.87 | 0 |

Data source: TFP estimated using capital stock data retrieved from the St. Louis Federal Reserve research database, labor data from The Conference Board, and GDP data from the World Bank.

The Difference-in-Differences equation runs:

$$y_{ist} = \gamma_s + \lambda_t + \delta * D_{st} + \epsilon_{ist}$$

(16.0)

The dependent variable in this test is the implied Total Factor Productivity levels. The $D$ dummy variable indicates whether an observation is affected by having undergone the treatment, here any post-1990 observation for Israel, being—as demonstrated in the above table—the Israeli entry for 1995 alone.

$$TFP_{ist} = \gamma_s + \lambda_t + \delta * Treatment_{st} + \epsilon_{ist} \tag{17.0}$$

We run the regression and find a very significant positive estimate of the coefficient to the treatment ($\hat{\delta} = 1.051$), implying that the treatment positively affected the productivity of Israel.

**Table 10. Differences-in-Differences estimation**

| VARIABLES | (1)<br>ImpliedTFP |
|---|---|
| 1.time | 0.141 |
| | (1.034) |
| 1.Treatment | $1.051^{***}$ |
| | (0) |
| Constant | 1.731 |
| | (1.034) |
| Observations | 4 |
| R-squared | 0.482 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Although the results are very statistically significant, it is difficult to say whether the effect was specifically due to changes in language composition. The large-scale immigration of Russian Jews to Israel could very plausibly have had many different effects on Israeli productivity, through—for instance—the influx of highly-educated workers. Isolating the language effect is very difficult, making the Difference-in-Differences results not very useful. The intention of this Difference-in-Differences regression is to illustrate the inherent difficulty in testing for indications of a causative relationship between linguistic fractionalization and productivity, as large relatively immediate changes in language composition tend to be the result of other factors with far-reaching effects, such as war, societal upheaval, large-scale immigration, and so forth – from which extracting the effect of language itself is very difficult. Further, constructing a traditional treatment study—rather than looking in history for something resembling a treatment event for the purposes of Differences-in-Differences—is extremely difficult on this level and in this context.

In order to better capture the effect of language on productivity, one would perhaps need to perform Difference-in-Differences wherein the treatment group is a region whose otherwise fairly static population undergoes significant language composition changes, due to perhaps a radical change in language policy. The difficulty of finding such a case, however, precludes it from common usage. In the absence of a viable test for giving support for either of the possible causative direction—let alone for any causative relationship to be present at all—we descend into what can ultimately be characterized as a speculative analysis.

Generally, that countries with higher levels of productivity attract greater quantities of migrants is fairly intuitive, but an acceptance of the plausibility of the idea of a causal relationship in the other direction—i.e. there being causality of linguistic fractionalization on productivity—does not necessarily preclude one from also, to at least some degree, accepting the other. A bi-directional effect could exist, with the language diversity itself having some direct causative effect—whether positive or negative—on a country's TFP, while simultaneously pre-existing high productivity levels to a large degree being what attracted the immigration that caused the fractionalization beyond the levels caused by historical distinct language communities within the country.

## 7.0 Conluding Comments

In a multicultural society where labor capital is not restricted by borders there are countless situations wherein people of different cultures and nationalities need to interact with each other. As language is at the core of communication, it is a pivotal factor in determining the efficiency of information transfer between two parties. This efficiency is ruled by a number of different mechanisms that take place between two people who speak different languages. Our objective was to study whether or not these mechanisms had enough of an impact on communicative efficiency to affect the overall productivity of economically successful nations.

Using existing theory on language economics we defined two working hypotheses. Firstly we expected to find a positive relationship between linguistic diversity and productivity, a bold idea as this is the opposite of what the majority of the literature hypothesizes. Secondly, we expected that the measure that would provide the highest level of significance would be a measure that included linguistic distance as well as diversity.

To examine this, five different measures of linguistic diversity were calculated for each of the 34 countries of the OECD and regressions of TFP on these measures were run. The results were able to confirm our hypotheses. We found that there was a positive relationship between linguistic diversity and TFP regardless of if we used a fractionalization or a polarization measure. We also found that the measure that provided the highest level of significance as well as the best fit was the ER index, which is a polarization measure that accounts for linguistic distances. The prominence of the latter attribute is consistent with our hypothesis. We did not expect however, that the measure that best explained the variance in TFP would be a polarization index. This type of measure is often used in relation to studies on civil conflict rather than productivity. It is especially surprising given the positive correlation that we found, since polarization more often is related to social antagonism, which one could imagine making coordination more difficult. One possible reason for why a polarization measure would provide such a result is that a small number of larger groups are more inclined to cooperate with each other internally in order to compete against and other groups, a competitiveness that could drive productivity and growth. We also did not expect that the other two measures that likewise included distances would not be able to provide similarly significant results. They performed even worse than the measures that did not include distances at all.

Our results differ from the common trend in language economics where most researchers find negative relationships between linguistic diversity and various economic variables. There are two main reasons as to why our results may diverge. The first is our choice of countries. Previous research has focused more on development studies or has set geographical delimitations. Few have focused on an economically and politically distinct set of countries as the OECD. A second reason for the differences described above is the manner in which we produced the linguistic diversity measures. We used a higher degree of standardization when accounting for group shares compared to previous research since the language data comes from many different sources with different levels of disaggregation. We have also treated linguae francae differently to fit the aim of the study and to avoid providing misleading results.

Future studies should focus on solving the issue of causality. To do this, language data over longer time spans is necessary, which is scarcely available at this time. It would also be interesting to complete this study with a larger set of countries to be able to fully compare the differences in the relationship between linguistic diversity and productivity between countries of different geographic areas, or countries at different levels of economic success. Another idea would be to redo the study with the inclusion of languages at different levels of disaggregation. For example, one might include dialects, as these can also affect the quality of communication—and perhaps one could thus provide a more nuanced view of the issue.

# References

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlay, S., and Wacziarg, R. (2003). "Fractionalization," *Journal of Economic Growth*, NY, 8(2), pp.155-94.

Alesina, A. and Ferrara, E. (2005). "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, Pittsburgh ,43(3), pp.762-800.

Ashraf, Q. and Galor, O. (2013), "The 'Out of Africa Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *American Economic Review*, Pittsburgh, 103(1), pp.1-46.

Bloomberg, (2015) "Bloomberg Innovation Index 2015." http://www.bloomberg.com/graphics/2015-innovative-countries/ (Accessed: May 2 2016).

Bond, N. and Ginsburgh, V. (2016). "Language and Emotion," *The Palgrave Handbook of Economics and Language*, Basingstoke, UK: Palgrave Macmillan UK. pp.231-260.

Boroditsky, L. (2011). "How Language Shapes Thought," *Scientific American*, NY, February 2011, pp.63.

Central Intelligence Agency, (2016). "Country Comparison: Exports," CIA World Factbook. https://www.cia.gov/library/publications/the-world-factbook/rankorder/2078rank.html (Accessed: April 26 2016).

Central Intelligence Agency, (2016). "South Korea," CIA World Factbook. https://www.cia.gov/library/publications/the-world-factbook/geos/ks.html (Accessed: April 21 2016).

Chen, K.M. (2013). "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets," *American Economic Review*, Pittsburgh, 103(2), pp.690-731.

Chiswick, B.R. and Miller, P.W. (1992). "The Determinants of Post-Immigration Investments in Education," Economics Discussion / Working Papers 92-17, University of Western Australia, Department of Economics.

Collier, P., (2001). "Ethnic diversity: An economic analysis," *Economic Policy: A European Forum*, Oxford, 32, pp.129-166.

Desmet, K., Weber, S., and Ortuño-Ortín, I. (2005). "Peripheral Diversity and Redistribution," CORE Discussion Paper, Lovain-la-Neuve, No. 44.

Desmet, K., Weber, S., and Ortuño-Ortín, I. (2009). "Linguistic Diversity and Redistribution," *Journal of the European Economic Association*, Cambridge, 7(6), pp.1291-1318.

Desmet, K., Ortuño-Ortín, I., and Wacziarg, R. (April 2014), "Culture, Identity and Diversity," Unpublished manuscript, UCLA.

Easterly, W. and Levine, R. (1997). "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, Hanover 112(4), pp.1203-1250.

Esteban, J. and Ray, D. (1994). "On the Measurement of Polarization," *Econometrica*, NY, 62(4), pp.819-51.

Fearon, J.D. (2003). "Ethnic and Cultural Diversity by Country," *Journal of Economic Growth*, NY 8(2), pp.195-222.

Florida, R. and Gates, G. (2001). "Technology and Tolerance: The Importance of Diversity to High- Technology Growth," The Brookings Institution – The Center of Urban and Metropolitan Policy, Washington, DC (June) pp. 1-12.

Florida, R., (2002). The Rise Of The Creative Class: And How It's Transforming Work, Leisure, Community And Everyday Life, NY, NY: Basic Books.

FRED Economic Data. Federal Reserve Bank of St. Louis. Web. January-May 2016.

Gay, V. Santacreu-Vasut, E. and Shoham, A. (2013). "The Grammatical Origins of Gender Roles," Berkeley Economic History Laboratory (BEHL) Working Paper Series, Berkeley, WP2013-03.

Ginsburg, G. and Weber, S. (Eds.) (2016). *The Palgrave Handbook of Economics and Language*, Basingstoke, UK: Palgrave Macmillan UK.

Grafton, R., Knowles, S. and Owen, P.D. (2004). "Total Factor Productivity, Per Capita Income and Social Divergence," *Economic Record*, Sydney, 80, pp.302-313.

Grafton, R., Kompas, T., and Owen, P. (2007). "Bridging the Barriers: Knowledge Connections, Productivity and Capital Accumulation," *Journal of Productivity Analysis*, NY, 28(3), pp.219-231.

Greenberg, J.H. (1956). "The Measurement of Linguistic Diversity," *Language*, Washington, DC, 32(1), pp.109-115.

Hall, R.E. and Jones, C.I. (1996). "The Productivity of Nations," National Bureau of Economic Research (NBER) Working Paper, Cambridge, MA, No. 5812.

Hulten, C.R. (2000). "Measuring Innovation in the New Economy," National Bureau of Economic Research (NBER) Working Paper, Cambridge, MA, No. 7471.

Israeli Ministry of Education, Culture and Sports, (2007). "עליית יהודי ברית המועצות," (Hebrew), (title translation: The immigration of the Jews of the Soviet Union). http://cms.education.gov.il/EducationCMS/Units/Tziyonut/AliyatBritHamoatzot.htm (Accessed: May 4 2016).

John, A., (2016). "Dynamic Models of Language Evolution: The Economic Perspective," *The Palgrave Handbook of Economics and Language*, UK: Palgrave Macmillan UK, Basingstoke. pp.101-120.

Jones, C.I., (1999), "Was an Industrial Revolution Inevitable? Economic Growth Over the Very Long Run," National Bureau of Economic Research (NBER) Working Papers, Cambridge, MA, No. 7375.

La Porta, R., Lopez de Silanes, F., Shleifer, A., and Vishny, R. (1999). "The Quality of Government," *Journal of Law, Economics and Organization*, Oxford, 15(1), pp. 222-279.

Levine, R. (2005). "Finance and Growth: Theory and Evidence," National Bureau of Economic Research (NBER) Working Paper, Cambridge, MA, No. 10766.

Mappes, T. and Puurtinen, M. (2009). "Between-group Competition and Human Cooperation," *Proceedings of the Royal Society B*, London, 276, pp.355-360.

Marschak, J. (1965). "Economics of Language," *Behavioral Science*, Dordrecht, 10(2), pp.135-140.

Mauro, P. (1995). "Corruption and Growth," *Quarterly Journal of Economics*, 110(3), Hanover, pp.681-712.

Mealem, Y. and Melnick, R. (2009). "Israel's Economy 1986-2008," Jewish Virtual Library Publications, September 2009. http://www.jewishvirtuallibrary.org/jsource/isdf/text/Melnick_Mealem1.html (Accessed: May 4 2016).

Montalvo, J.G. and Reynal-Querol, M. (2005), "Ethnic Polarization, Potential Conflict, and Civil Wars," *European Journal of Political Economy*, Amsterdam, 21, pp.445-465.

Narodav Mira Geo-referencing of ethnic groups (GREG). Weidmann, B.N., Rød, J.K., and Cederman, L.E. (Eds.) (2010). Web. January 2016.

OECD.stat. Organisation for Economic Co-Operation and Development (OECD). Web. January-May 2016.

Output, Labor and Productivity 1950-2015. The Conference Board Total Economy Database, Sept. 2015. Web. January 2016.

Parotta, P., Pozzoli, D., and Sala, D. (2014). "Ethnic Diversity and Firms' Export Behavior," IZA Discussion Papers 7923, Institute for the Study of Labor (IZA), Bonn.

*Religionen & Weltanschauungsgemeinschaften in Deutschland: Mitgliederzahlen*, Religionswissenschaftlicher medien- und informationsdienst E.v (REMID), (German). http://remid.de/info_zahlen/ (Accessed: March 10 2016).

Reynal-Querol, M. (2002). "Ethnicity, Political Systems, and Civil Wars," *Journal of Conflict Resolution*, Thousand Oaks 46(1), pp.29-54.

Santacreu-Vasut, E., Shenkar, O. and Shoham, A. (2014). ”Linguistic Gender Marking and its International Business Ramifications,” *Journal of International Business Studies*, Basingstoke, 45, pp.1170-1178.

Spolaore, E. and Wacziarg, R. (2015). “Ancestry, Language and Culture,” CAGE Working Paper Series, Coventry, No. 234.

Tinbergen, J. (1962). Shaping the World Economy: Suggestions for an International Economic Policy, New York, NY: The Twentieth Century Fund.

World Bank Open Data. The World Bank. Web. January-May 2016.

# Appendices

## Appendix A – Variable List

| | |
|---|---|
| *ImpTFP* | Implied Total Factor Productivity. Calculated by the authors as described in the Methodology section. *lImpTFP* is its logarithm. *Source: Data for calculations retrieved from the World Bank and the OECD statistical databases.* |
| *Ling_frac* | Linguistic Fractionalization index—between 0.00 and 1.00—based on the ELF index. Values for sample calculated by the authors. *lLing_frac* is its logarithm. *Data sources for language compositions used to calculate indices: Alesina 2003, CIA World Factbook and Ethnologue.* |
| *RQ* | Reynal Querol index—between 0.00 and 1.00—from Reynal-Querol (2002). Values for sample calculated by the authors. *lRQ* is its logarithm. |
| *GI* | Greenberg index—between 0.00 and 1.000—from Greenberg (1956). Values for sample calculated by the authors. *lGI* is its logarithm. |
| *ER* | Esteban-Ray index—between 0.00 and 1.00—from Esteban and Ray (1994). Values for sample calculated by the authors. *lER* is its logarithm. |
| *PH* | Peripheral Heterogeneity index—between 0.00 and 1.00—from Desmet, Weber and Ortuño-Ortín, (2005). *lPH* is its logarithm. |
| *Rel_frac* | Religious fractionalizaton index—between 0.00 and 1.00—from Alesina. *lRel_frac* is its logarithm. *Source: Alesina 2003.* |
| *Ethnic_frac* | Ethnic fractionalization index—between 0.00 and 1.00. *lEthnic_frac* its logarithm. *Source: Alesina 2003.* |
| *Yrseduca* | Nationally average years of education. *lYrseduca* is its logarithm. *Source: OECD database.* |
| *_POP* | Total national population. *lPOP* is its logarithm. *Source: World Bank.* |
| *Catholic* | Dummy variable for historically majority-Catholic countries. *Source: Religious composition data retrieved from CIA World Factbook, Pew Research, and REMID.* |
| *Protestant* | Dummy variable for historically majority-Protestant countries. |
| *Islam* | Dummy variable for historically-majority Islamic countries |
| *Judaism* | Dummy variable for historically majority-Jewish countries. |
| *Englishlegalorigin* | Dummy variable for countries with a predominantly English legal origin. *Source: Legal origin data retrieved from La Porta 1998, CIA World Factbook.* |
| *Frenchlegalorigin* | Dummy variable for countries with a predominantly French legal origin. |

| | |
|---|---|
| *Socialistlegalorigin* | Dummy variable for countries with a predominantly Socialist legal origin. |
| *Scandinavianlegalorigin* | Dummy variable for countries with a predominantly Scandinavian legal origin. |