

1 2 9 0



UNIVERSIDADE D
COIMBRA

André de Abreu Saraiva Monteiro Alves

**ADMINISTRATIVE DISTANCE IN
INTERNATIONAL BUSINESS STUDIES: AN
EXTENSION PROPOSAL.**

**Dissertação no âmbito do Mestrado em Gestão orientada pelo
Professor Doutor Fernando Manuel Pereira de Oliveira Carvalho e
apresentada à Faculdade de Economia da Universidade de Coimbra.**

Junho de 2020



FEUC FACULDADE DE ECONOMIA
UNIVERSIDADE DE COIMBRA

André de Abreu Saraiva Monteiro Alves

Administrative Distance in International Business studies: An extension proposal.

Dissertação de Mestrado em Gestão, apresentada à Faculdade de Economia da Universidade de Coimbra
para a obtenção do grau de Mestre

Orientador: Professor Doutor Fernando Manuel Pereira de Oliveira Carvalho

Coimbra, Junho de 2020

Agradecimentos

Nesta fase final deste desafio, há várias pessoas a quem tenho uma enorme gratidão e que gostaria de agradecer.

Ao Professor Doutor Fernando Manuel Pereira de Oliveira Carvalho, que me orientou neste ano, pela sua disponibilidade, interesse, boa disposição e conselhos. Não só me desafiou, como teve enorme paciência perante todas as minhas questões. Se me deu liberdade para explorar tópicos que me interessassem, nunca deixou que me perdesse.

Ao Professor Doutor Marcelo Pereira Duarte, que não só sempre me ajudou quando não percebia algum detalhe técnico, como me permitiu ter conversas alongadas acerca dos diversos tópicos neste trabalho, em que várias ideias e conceitos foram cristalizados.

Finalmente, gostaria de agradecer à minha família. Não só me suportaram durante todo este percurso, como se certificaram que apanhava algum sol, quando me esquecia durante demasiado tempo.

Epigraph

“It’s the sides of the mountain which sustain life, not the top. Here’s where things grow. But of course, without the top you can’t have any sides. It’s the top that defines the sides. So on
we go...”

Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance*

Abstract

The concept of Distance as a metaphor for the more descriptive concept of hardship to traverse, and its operationalization are of unquestionable topicality and relevance, being used extensively in scientific literature to study diverse phenomena in International Business. Despite the existence of diverse models presented in different forms, all have limitations. In this dissertation we will focus on the Institutional Distance model proposed by Berry, Guillén and Zhou, presented in 2010, to study in particularity one of its components.

Studies that narrow on a more specific aspect of the Institutional Distance and study in with more depth are relatively less frequent. The objective of this dissertation is to develop said construct of Administrative Distance, which we do with the proposal of an extension of concept, operationalizes with the addition of a Vertical Administrative Distance variable, calculated using the *Doing Business Report* data from the World Bank. A fundamental property of the *Doing Business Report* for this choice is its reliance on more explicit bureaucratic matters.

Methodologically we will use as a reference model to test our proposal the study made by Duarte and Carvalho in 2018 relating for Foreign Direct Investment in Portugal, using Mahalanobis distances and panel regression models.

We believe that the results obtained indicate that our proposal of extension of Administrative Distance is beneficial because its efficacy in measuring what it attempt to is improved, as well as its consistency. This work also opens several avenues for future research.

Keywords: Administrative Distance, Institutional Distance, Doing Business Report, Panel Data, Mahalanobis Distance

Resumo

O conceito de Distância como metáfora para um conceito mais descritivo de dificuldade de ultrapassar, e a sua operacionalização são de uma actualidade e relevância inquestionáveis e têm sido extensivamente usados na literatura científica para estudar diversos fenómenos em Negócios Internacionais. Apesar de existirem vários modelos apresentados em diferentes formas, todos apresentam limitações. Nesta dissertação vamos-nos focar no modelo de Distância Institucional proposto por Berry, Guillén and Zhou apresentado em 2010 para estudar de forma particular uma das suas componentes.

Estudos que se focam num aspeto mais específico de Distância Institucional e o estudam de forma mais profunda são relativamente pouco frequentes. O objetivo desta dissertação é desenvolver o construto de Distância Administrativa, apresentando uma proposta de extensão do conceito, operacionalizada com a adição de uma variável de Distância Administrativa Vertical calculada usando os dados do Doing Business Report do World Bank. Uma propriedade fundamental dos dados do Doing Business Report para esta escolha é o seu maior foco em matérias explicitamente burocráticas.

Metodologicamente vamos utilizar como modelo de referência para testar a nossa proposta o estudo feito por Duarte e Carvalho em 2018 sobre Investimento Direto Estrangeiro em Portugal, utilizando distâncias Mahalanobis e modelos de regressão de dados em painel

Acreditamos que os resultados obtidos indicam que a nossa proposta de extensão da Distância Administrativa é benéfica porque melhora a sua eficácia a medir aquilo a que se propõe, assim como melhora a sua consistência. Este trabalho permite também a abertura de várias portas para pesquisa futura.

Palavras Chave: Distância Administrativa, Distância Institucional, Doing Business Report, Dados em Painel, Distância Mahalanobis

Table of Figures

- Table 1 – Regression without Vertical Administrative Distance 54
- Table 2 – Regression with Vertical Administrative Distance..... 56
- Table 3 – Correlation Matrix with Vertical Administrative Distance 57
- Table 4 – Correlation Matrix of FDI and Vertical Administrative Distance Pillar Variables 59
- Table 5 – Isolated Vertical Administrative Distance Pillar Variables..... 59
- Table 6 – Reduced Model with Vertical Administrative Distance Pillar Variables 60
- Table 7 – Normalized Regression with Vertical Administrative Distance 62
- Table 8 – Normalized Isolated Vertical Administrative Distance Pillar Variables 63
- Table 9 – Integrated Administrative Distance Regression 64
- Table 10 – Baseline Study Model 1 – Summary Statistics..... 81
- Table 11 – Baseline Study Model 1 – Correlation Matrix..... 81
- Table 12 – Baseline Study Model 1a - GLS Regression..... 82
- Table 13 – Baseline Study Model 1b - GLS Regression..... 82
- Table 14 – Baseline Study Model 2 – Summary Statistics..... 83
- Table 15 – Baseline Study Model 2 – Correlation Matrix..... 83
- Table 16 – Baseline Study Model 2a - GLS Regression..... 84
- Table 17 – Baseline Study Model 2b - GLS Regression..... 84
- Table 18 – Model 1, with Vertical Administrative Distance addition – Summary Statistics 85
- Table 19 – Reduced Model with Vertical Administrative Distance addition..... 85
- Table 20 – Model 5, with Starting a Business variables, not normalized..... 86
- Table 21 – Model 6, with Dealing with Construction Permits variables, not normalized..... 87

Table of Contents

1 – Introduction.....	10
2 – Literature Review	13
2.1 – The Underlying Objective of the Metaphor of Distance as Hardship to Traverse	13
2.2 – Illusions in Distance	16
2.3 – Descriptor and Underlying Concept Issues	20
2.4 – Aggregation within Reflective Constructs	22
2.5 – Horizontality and Verticality.....	23
2.6 – Conceptualization of Institutions	24
2.7 – The evolution of Distance Measures.....	29
2.8 – Institutional Distance Models	31
2.9 – Administrative Distance as a component of Institutional Distance Models and its Conceptualization.....	34
3 – Study Proposal for the Improvement of Administrative Distance	38
3.1 – Hypothesis Formulation	42
4 – Methodology	44
4.1 – Dependent Variable	44
4.2 – Independent Variables	45
5 –Results	52
5.1 – Baseline Study	52
5.2 – Vertical Administrative Distance Model.....	53
5.3 – Integrated Administrative Distance Preliminary Model.....	63
6 – Discussion and Conclusion	66
6.1 – Contribution	67
6.2 – Limitations.....	68
6.3 – Avenues for Future Research	69
5 – Bibliographic References	71
6 – Appendix.....	81

1 – Introduction

In this dissertation, we will attempt to further develop the construct and measurement of Administrative Distance, with aim to improve its efficacy in measuring what it attempts to and improving its consistency. Other studies have shown Administrative Distance to be of particular relevance, both empirically and theoretically (Ghemawat, 2001; Berry, Guillén and Zhou, 2010; Ambos and Håkanson, 2014; Reimann, Rauer and Kaufmann, 2015; Ghemawat and Hout, 2017; Kang, Lee and Ghauri, 2017; Quer, Claver and Rienda, 2017; Beugelsdijk, Ambos and Nell, 2018; Knoll *et al.*, 2018; Liu *et al.*, 2019; Belderbos, Du and Slangen, 2020), despite not being common that it is the focal point of research.

This topic of investigation onto distances in International Business, with particular focus on Administrative Distance, has high relevance for the ubiquity its usage. Distances, in International Business, are the yardstick that aims at measuring the nature or hardship that certain international interactions, or interactions with the “other”, will have (Zaheer, Schomaker and Nachum, 2012). Therefore, it is in its nature central in conducting research in the field of International Business, or rather, Administrative Distance holds itself as an important distance, belonging to a larger group of distances, which are in themselves central. Nevertheless, there are considerable developments that have been identified as having to be made both in distances in general, as well as in Institutional Distance (Shenkar, 2001; Zaheer, Schomaker and Nachum, 2012; Hutzschenreuter, Kleindienst and Lange, 2016; Beugelsdijk, Ambos and Nell, 2018), which we will apply in a more focused manner to Administrative Distance.

For these purposes, in the literature review we will start by moving gradually through the conceptualization of distance, highlighting its metaphorical nature and underlying objective of measuring the hardship of transversal. We then continue, still focused on the clarification of distance as concept, highlighting several of the illusions that have been identified in it, and particularly in contextual distances. The next chapter, shifts focus away from distance as a stand alone concept but highlight further issues that exist, however this time focused on the lack of correspondence between descriptors, or the names given to things, and the underlying

concepts, that is, the things themselves. Onto the fourth subsection of the literature review we will now develop on the considerations that should be had concerning aggregation, within distance constructs, particularly again in reflective ones. Lastly, in this vein of the literature review we evaluate the differentiation between measures with horizontal or vertical characteristics, being that all these subchapters serve our discussion later on, providing both technical vocabulary and conceptual clarity. The focal point changes slightly when we will attempt in the seventh subchapter to achieve conceptual clarity with regards to institutions, being that the eighth chapter follows the same topic but onto Institutional Distance models. Lastly, within all of the literature review we come to our focal topic, with the benefits of contextualization had beforehand, where we approach Administrative Distance within Institutional Distance models, as well as its conceptualization. Due to the nature of both Institutional Distance, and Administrative Distance, we believe that the choice of having two separated chapters for Institutional Distance, while an integrated chapter for Administrative Distance will be more conducive to clarity and an appropriate flow of interconnected information. The first five subchapters, then, are concerned with providing the tools for discussion, the sixth through eight with the needed contextualization, and the ninth with the review of the central topic at hand: Administrative Distance.

In the third section, we deal with the study proposal and hypothesis formulation, where we describe in detail what we are going to do and test, as well as our reasons to do so. Here we identify the primary issues surrounding Administrative Distance within Berry, Guillén and Zhou's (2010) Institutional Distance model, which we will use, as well as how our proposed addition might be able to limit said issues, and potentially improve the Institutional Distance model itself.

In the fourth section of this work we will start dealing with Administrative Distance models proper. After presenting both dependent and independent variables, we will use a baseline study to analyze Administrative Distance, being that we will be using Foreign Direct Investment (FDI) as an initial testing application of the model and its capacity, as well as Portugal for the initial location of testing, by virtue of being the topic of choice of said baseline study. The dissertation, itself, is not focused on FDI proper, or Portugal, but simply uses them as a one

use case, like many possible others, to test any proposed framework of Administrative Distance. This, in turn, makes it hard not to speak at least in passing about FDI and Portugal, which we will do when necessary.

After looking our baseline study, we will transform it as necessary for the possible addition of any new variables, and test said addition. This is done to, in effect, be able to have a measure to access if the model is in some way improved. In this section, we will also look more deeply onto the variables present in the *Doing Business Report*. Lastly, we will briefly look into the possibility of creating an Integrated Administrative Distance, since up until now both stood as separated variables, despite theoretically and conceptually connected.

At the end of this work, we expect either one or two things. Either a better measure of Administrative Distance can be proposed, or if not, then we will have at least understood better the behavior of this construct, as well as its limitations.

2 – Literature Review

2.1 – The Underlying Objective of the Metaphor of Distance as Hardship to Traverse

Right at the early stages of this work, it might be beneficial to clarify what we mean by distance, in the context of International Business. Often times this point is foregone, as though it is self-evident, particularly when distance is not the topic at hand, but is just being used instrumentally. We make the case that it is not so, as the types of distances we are treating are conceptually quite different from the type of distances found in mathematics, or even concrete physical distances (Shenkar, 2001). We posit that a key point necessary to the discussion and development of such a construct, is the conceptual clarity that we are speaking of a metaphor (Shenkar, 2001; Zaheer, Schomaker and Nachum, 2012; Ambos and Håkanson, 2014).

While the previous statement is self-evident, getting through to underneath said metaphor might prove useful later on, and it might be worthwhile to make it explicit and descriptive. Hutzschenreuter, Kleindienst and Lange (2016) round down the possible types of distances to six: cultural, institutional (in which, we believe, administrative distance is implicitly included), geographic, economic, psychic and other. With some exceptions like Shenkar (2001, 2012), Zaheer, Schomaker and Nachum (2012) attempts at clarification of the concept itself seem relatively scarce, particularly in a situation where as Hutzschenreuter, Kleindienst and Lange (2016) put it: “Given four decades of distance-related research one would expect a consensus on the conceptualization of distance would have been reached. And yet, there is still ambiguity...”

One of the reasons underlying this has analogous traits to the “psychic distance paradox” (O’Grady and Lane, 1996) , where decision makers which consider a foreign market as close, tend to be unobservant of and unprepared for differences, big and small, which might then interfere with the adequacy of the decision in question. If we can trace the parallel, such a

thing as the concept of distance in International Business might appear so basic as to not warrant further discussion.

However, even in its most “natural” form, Geographical Distance, which we will approach first as it is the candidate least likely to hold substantial differences from the abstraction of mathematical distances, what is being considered is not even then the distance itself, but rather the hardship of traversing that distance. Arguably, no one in this context is concerned with the *actual* physical space between countries. Meaning, if from our metaphorical village there is another, to the north, through which we have to traverse pleasant meadows, and the same distance, to the south, that instead has a harsh desert, despite both being the same Geographical Distance from us, they are not identical for the purposes of establishing a commercial route. If infrastructure is built, “distances” can shrink. If particularly harsh geographical attributes exist, like a desert or a mountain, they can increase as compared to a generic baseline. If we are speaking geographically, however, distances will generally be about symmetrical, discounting elevation changes or tides and winds, and not positive since a cost will have to be incurred to transit them.

However, regarding geographical distances, frequently just some form of the literal distance is measured. Malhotra, Sivakumar and Zhu (2009), as an example, use the distance in kilometers between capitals, justifying it as it has been used in previous studies. Specifically the authors reference Buckley *et al.* (2007) and Ojala and Tyrväinen (2007). Buckley *et al.* (2007) on their part, do not lose too much time on the geographical component, as it is a mere control to isolate geographical effects from other variables, justifying only that it should have a negative relation to FDI. Ojala and Tyrväinen (2007) spend more time handling the theory behind why Geographical Distance matters. The authors in fact do not use the capital, as the previous studies, but the center of mass of the countries at hand. However, in the methodology section it is not approached, also, why this is a valid simplification. Berry, Guillén and Zhou (2010), for their part, use the “Great Circle Distance”. At the onset, it is not evident that it will even *be* a problematic simplification, given that many studies focus on effects at country level, nevertheless seems it is not a simplification that is often addressed. Belderbos,

Du and Slangen (2020), as another example, point to the existence of significant intra-country heterogeneity, through the study of the differentiated aspects of global cities, which legitimizes such questions.

Some simplification of geographical distance is not inherently wrong even under this conceptualization, and even if the effects studied are not at the country level. It just leaves the implicit assumption that geographical factors will, approximately, average out; the quality of infrastructure will, approximately, average out; and some, potentially, non-significant precision will be lost. Generally, then, Geographical Distance is also stable in time. Directly that might again be so obvious as to not warrant discussion.

Following this logic, that is not universal in nature, but rather used as an example of concept, we can say that Portugal is just as relatively far away from Russia as it was a thousand years ago. Not because the literal geographical distance is the same, but because all the countries have gotten closer together geographically, at the approximately same rate, due to the role of infrastructure and technology development. Then, the literal usage of distance can be adequate for comparisons in certain situations, given that we accept forgoing some precision, and particularly if we are studying effects at the country level, particularly between countries with relatively homogenous infrastructure. On the other hand, seeing the major efforts some countries are putting into developing connecting infrastructure across the globe, (e.g. China's Belt and Road) it is not self-evident that these effects average out, or are statistically insignificant, at least conceptually, although evidently such hypothesis could only be confirmed or rejected through empiric studies. Our reason for addressing it, in short, was to simply highlight that it does not have a *self-evident* nature.

When we talk about constructs such as Cultural Distance, then, it is of course a metaphor. We can "walk" from a culture to another, but these are not literal meanings. However, reiterating this aspect again, we are not actually concerned with the distance itself, but rather with the "hardship to traverse" implicit in it. Beugelsdijk, Ambos and Nell (2018) further divide distances between geographical distances, and contextual distances – cultural, economic and administrative, since the authors are using the CAGE framework (Ghemawat, 2001) as a

guideline. The authors further this distinction by pointing out certain properties of geographical distances (like symmetry) do not extend to contextual ones. Here we started by making the case that even Geographical Distance, in this context, has certain contextual attributes, regardless of their significant impact in generalized practice, in order to support the case that contextual differences, then, are likely to have even further distinctive behaviors from the two points in two dimensions mathematical abstraction.

The proposition of Shenkar (2001, 2012) to use friction in Cultural Distance measures highlights just the same aspect as distance does, even if the metaphor can be more flexible in nature. Under the concern in “friction” is the implicit concern with the “hardship to move”, whether through drag, or slippage, meaning too much, or too little, friction.

The same reasoning applies for administrative or other kinds of hardships or distances companies might face. However the default metaphor will remain the one of “distance”, as it is what we will nevertheless use. Regardless, the consideration that by it, or friction, what is meant, explicitly or implicitly, is a metaphor for the more descriptive *hardship to traverse, deal, work with, interact*, or identical.

2.2 – Illusions in Distance

Having in the previous chapter approached the issue of conceptual clarity more broadly, we believe it is of importance to also now consider the illusions that exist within the current conceptualization of distance in more depth (Shenkar, 2001). Shenkar (2001, 2012) has greatly influenced the field of not only Cultural Distance, but also the expanded field of this type of contextual distance measurement and construct development (Beugelsdijk, Ambos and Nell, 2018; Dow, 2018).

Shenkar (2001) analyzes the conceptualization and methodology associated with Cultural Distance constructs. However many of them are directly or almost directly transferable between Cultural Distance and Administrative Distance. This is, the arguments made for their applicability in Cultural Distance constructs seem to generally hold true for other types contextual distances (Beugelsdijk, Ambos and Nell, 2018; Dow, 2018). The author goes into

considerations that should be had during construct development, proposes friction as a better metaphor and leaves recommendations for the development Cultural Distance measures, being that he is using the Kogut and Slight (1988) index as his benchmark for these observations. This becomes particularly relevant, as precisely what we are attempting to develop is the construct and measurement of Administrative Distance.

Dealing with conceptual properties, Shenkar (2001) points out five illusions regarding conceptual properties that, at the very least, can be challenged regarding the cultural distance frameworks.

The first is the illusion of symmetry, where it is challenged that the distance from A to B might not be the same as from B to A. We have touched upon this in the first chapter of this literature review, which leads to direction being an important attribute of Distance (as an example: Qian and Sandoval-Hernandez, 2016; Konara and Shirodkar, 2018; Mueller, Hendriks and Slangen, 2018). As we have framed before, even in literal distances, the “work needed to traverse”, which we contended is a clearer descriptor for the underlying objective of these metaphors, is not necessarily the same. While the Mount Everest’s top is at the same distance from the bottom, as the bottom is from the top, the journeys are not alike.

The next illusion is that of stability and it poses the questions: do cultures and administrations never change, or even, do they change but do so in tandem? If they change, what is the rate at which they have been changing, and in which direction? Do we observe convergence or divergence? (Shenkar, 2001) This illusion is more straightforward in nature and simply warns that the assumption that things do not change, while might be reasonably valid in the short term, where things change little, on the medium and long term might become quite invalid, such as if compounding effects start to set in and become relevant. This illusion also makes reference to the level of analysis which is to be undertaken. While at an aggregate level time might have effects because of institutional or cultural developments, at the individual level the preponderance of time effects might have to do more with learning and network formation (Johanson and Vahlne, 1977, 2009).

The next illusion posed by Shenkar (2001) is that of linearity. At its core, the call of this illusion is relative to the assumption that, whatever metric we are using, something twice as distant is twice as foreign. Likewise, if we incorporate learning effects, having double the time of interaction to learn from it, will wield twice the knowledge of it. Shenkar (2001) calls to expatriate literature to exemplify that this might not be the case. Expatriates do not adapt in a linear fashion, but in a semi U-curve. In the beginning, although not in an optimal state, there is little abrasion in the foreign. However, as time moves, new differences are found, the expatriate recedes to his default state and suddenly integration and efficiency come to an all-time low. Only after this stage, the expatriate truly integrates into the foreign culture and reach the absolute, and not only local, optimum. In a more institutional setting, we can think of a company that in the beginning might only do small projects that bypass many of these challenges, and have only select human resources attributed to them (López-Duarte, Vidal-Suárez and González-Díaz, 2016; Kawai and Chung, 2019).

Shenkar's (2001) fourth illusion is that of causality. In it, specifically, Shenkar (2001) calls into question the assumption that cultural distance is the only relevant variable regarding Foreign Direct Investment (FDI), since it is the topic of analysis in his article. However, we can expand it to the analysis on anything. As long as we are only using one variable to analyze something, it is left implicit that that is the only relevant variable, and if not, that we are content to capture at least a partially spurious relationship using that variable, or have methodologically corrected for the omission of relevant variables¹. Even then, the explanatory power of any model in these circumstances will be much more limited (Wooldridge, 2013).

As this work pertains to the development of Administrative Distance, and in it a seemingly lack of substance despite relatively high levels of statistical significance being frequently found (Kuo and Fang, 2009; Blanc-Brude *et al.*, 2014; Quer, Claver and Rienda, 2017; Liu *et al.*, 2019), the issue of the causal mechanism of Administrative Distance is of particular relevance, as it influences the interpretation that can be made out of the results obtained. Additionally, if there is an omission of relevant variables, then their inclusion might either cause a drop in

¹ With some or other, possibly second best, control, as an example.

significance of the variables we had before, as the effects they previously had are now captured elsewhere, or an increase, as the effects they had are now better contextualized, or disentangled from the previously omitted relevant variables.

While this later effect seems at first sight less common, to get an empiric sense of it we might only imagine the likely outcome of trying to measure a sportsman ability to score by his or her salary. We might even find that as the salary increases the ability to score, as measured by salary, does not significantly increase. If this were to happen then, we would immediately think something wrong with our model and realize that we were not considering the league at which said sportsman was playing, or his or her position in the game. In this case, then, the inclusion of these omitted variable could possibly increase the statistical significance of salary as a predictor of scoring ability (Wooldridge, 2013, p. 90).² It is also just as possible that this illusion may play a role in less than ideal proxy variables. Even in ideal proxy variables, that little alter effectively the results obtained, it is worth the consideration over the nature which interpretation might take if we are not cognizant of this aspect present in a model³ (Wooldridge, 2013, p. 309)

The last illusion is that of discordance (Shenkar, 2001) and it goes into the assumption that Distance (or Friction) is universally bad. If we take, for the deconstruction of this illusion, the more flexible metaphor of friction, we can postulate the existence of complimentary attributes. That is, attributes that create a situation of friction more advantageous than the internal friction present in home culture. Even if we hold the construct of distance for the analysis of this illusion, to point that there is *only* discordance is only to hold transaction costs (Williamson, 1985) as relevant, while forgoing any learning or resources that might be acquired in the process of interacting through said distance (Johanson and Vahlne, 1977). This, at the very least, then, merges both contradictory effects, forgoing any possible benefits in clarity due to their de-interlacing. This exclusivity of transaction costs has been shown not to be the

² To clarify, the example given is not based on any study, nor are we saying it is the case. It exists simply to improve the conceptual clarity of a possible occurrence at the inclusion of a relevant variable, previously omitted.

³ (i.e.) Interpreting the proxy as opposed to the underlying variable and phenomena.

case (Malhotra, Sivakumar and Zhu, 2009; Verbeke, Puck and Tulder, 2018; Vljaj *et al.*, 2019), being that there is solid ground to say that, even if the aggregate effect is negative, there is room for the existence of positive effects. Additionally, there is evidence that a correct assessment of the true nature of distances is beneficial to the performance of companies (O’Grady and Lane, 1996; Azar and Drogendijk, 2018), which furthers the importance of this illusion.

2.3 – Descriptor and Underlying Concept Issues

Another relevant question to address, still at the early stages of this dissertation, is in regards to the descriptor of the type of distance we are talking about. As Beugelsdijk, Ambos and Nell (2018) have shown, the correlations between the results of Institutional Distance as measured by the Worldwide Governance Index (WGI), or Economic Freedom Index (EFI), while not as low as the similar indexes for Cultural Distance, are still quite low. Meaning, depending on what database we use, even before consideration of different methods, a different, uncorrelated result, is obtained. This raises questions about the validity of calling it all *Institutional Distance*, or yet, the results passing for *the* Institutional Distance. Differing results coming from different databases, that should, in concept, be identical, in the sense that they are measuring the same thing, not only brings forward the requirement, as the authors state, that the researcher duly justifies the choice, but also calls into question that it all is *the*, or even *the same*, institutional distance.

In order to interpret its results one has to be aware of both the methods of construction and the specific aspects of culture/administration/institutions that are measured, for which the simple name “cultural” or “institutional” does not give clue. This highlights the issue with overreliance on overarching concepts and simplified descriptors without proper support given to the underlying concepts these models and datasets explore, as well as the issue with an absence of explicit descriptors of these (Beugelsdijk, Ambos and Nell, 2018).

Moreover, as Hutzschenreuter, Kleindienst and Lange (2016) point out, there is not a consensus on the front of institutional distance. As they put it:

“Across the articles, the number, the type and the operationalizations of intrinsic characteristics used to proxy institutional distance differ greatly. For instance, regarding the measurement, some use the World Bank’s governance indicators (Gallego and Casillas 2014), others draw on the Economic Freedom Index (De Beule et al. 2014), yet others develop their own items (Chiao et al. 2010). Moreover, different authors have used identical measures for conceptually different intrinsic characteristics. This may be a reflection of a lack of agreement on the conceptualization of institutional distance.”

The authors then go on giving examples about how other articles have used “formal institutional distance”, “administrative distance” and “regulatory distance” all to refer to the differences in regulatory systems, while using the same WGI index. We present the following examples as complementing that:

- To Reimann, Rauer and Kaufmann (2015), Belderbos, Du and Slangen, (2020) and Malhotra, Sivakumar and Zhu (2009) *all* of the WGI indicators measure Administrative Distance;
- To Abdi and Aulakh (2012), *two* of the WGI indicators measure Formal Institutional Distance (North, 1991);
- To Dikova (2009) *all* of the WGI indicators measure Formal Institutional Distance,
- To Ang, Benischke and Doh (2015) *all* of the WGI indicators measure Regulatory Distance (Scott, 2014);

We are not, with this, implying that any particular author, in this case, is wrong in their use of the WGI. Rather that there is inconsistency and this inconsistency is relevant (Beugelsdijk, Ambos and Nell, 2018). We have no reason to believe this phenomenon is localized, and as we have stated, the above articles merely serve to provide a concrete example of the issue.

The previous paragraphs present a twofold problem. On one side a lack of a matching descriptor, and on the other the lack of a matching underlying concept, between different studies. The descriptors, that is, “Formal Institutional Distance”, “Administrative Distance”, “Regulatory Distance”, “Institutional Distance”, are used somewhat interchangeably, referring to possibly different underlying concepts captured by differing databases and indexes. Meaning, depending on which method/variables we use, we can achieve quite uncorrelated results, or quite weakly correlated.

Developing this to its conclusion, the literature review points to the fact that we can have both:

- Two studies about “Institutional Distance” that are quite different because one uses EFI, and the other WGI;
- Two studies treating “Formal Institutional Distance” and “Administrative Distance” that are really about the same thing because they use the same methods and databases.

These two issues relate to the inconsistency in the use of the same descriptor and inconsistency in the underlying concept of what said descriptor is referent to.

Beugelsdijk, Ambos and Nell (2018) highlight in this manner the non-existence of a general distance theory, being that the assumptions, mechanisms, and boundary conditions that are used will adapt based on the purpose of the distance being measured, however without a consistent framework guiding these decisions.

This creates the need that when choices must be made regarding a particular study, then proper justification and theorizing must be put in place to support these decisions, at the very least, due to the heterogeneity of models. This, in a manner beyond what would have to be done when using a model in an instrumental manner when all is quite set and defined. The ideal, however, would be to have a well-defined framework that allows for this interaction in a set manner (Beugelsdijk, Ambos and Nell, 2018).

2.4 – Aggregation within Reflective Constructs

Beugelsdijk, Ambos and Nell (2018) point out the choice between an aggregate construct of different things, or a single measure of one phenomenon. Some detail is lost in the building of the construct, and often conceptually different things might be aggregated, but the usefulness of the aggregation might prove worthwhile the loss, being that this should be evaluated on a case-by-case basis, as it seems very much to be the line of reasoning of Ghemawat (2001). In this case, it is desirable then that at least the direction of all component variables is generally the same, with the risk of containing contradictory effects that nullify each other (Shenkar, 2001).

In particular relevance towards this dissertation the authors make the conceptual distinction between a *formative* and a *reflective* construct. In the first, the components of the construct, each individually, built up to represent the underlying phenomenon. A reflective construct, on the other hand, has each of the components reflecting the underlying phenomenon, being that all of them put together will yield an improved, composite image of said phenomenon (Beugelsdijk, Ambos and Nell, 2018). To give concrete examples, GDP can be thought of as formative, while, to not use the example of the central topic of this dissertation, we can think of composite photography for the later.

Given that contextual distances are ostensibly non-directly observable, this extending onto distances applied onto them, they will always have a good measure of reflective characteristics. For example, looking at the WGI variables (Kaufmann, Kraay and Mastruzzi, 2011), its reflective nature becomes apparent, being that each of the components represents a latent facet of the institutional environment⁴. Nevertheless, this is an important conceptual property that should be taken into account in the building of a construct, particularly because of the still relatively soft ground within which contextual constructs, based on reflective measures, stand.

2.5 – Horizontality and Verticality

A useful distinction to make, when it comes to the variables used within distance constructs, is that of horizontality and verticality (Ghemawat and Hout, 2017).

To take the KS Index (Kogut and Singh, 1988) as an example, as well as the default equal weight regarding the different dimensions of culture it measures from (Hofstede, 1980), it does not matter if a country differs from another in the same direction on all four dimensions of culture, or if it alternates between directions in the deviation. This, since a lack of directionality, is characteristic of *horizontal measures of distance* (Ghemawat and Hout, 2017). Horizontal measures of distance, here in the sense that it is not attributable which direction is “better” or “worse”. In this sense, then, the study of directionality will have different characteristics

⁴ Them being: Voice and Accountability, Political Stability and Absence of Violence/Terrorism, Government Effectiveness, Regulatory Quality, Rule of Law and Control of Corruption.

than if we are handling a *vertical measure of distance*, in which a qualitative judgement in present (e.g. corruption levels) (Ghemawat and Hout, 2017).

Whether a country holds all high scores in Hofstede's (1980) dimensions, or, on the other hand, it scores all zeros is not indicative of a "good" culture or "bad" culture. This makes it so that the directionality of each of Hofstede's (1980) dimensions (i.e. is masculinity is represented by 100 or by 0) is irrelevant for the calculus of the KS Index, which is concerned solely with the difference between that of Country A and B.

With horizontal measures of distance, we may still study the differing effects of distance between countries. That is, the possible difference between effects of a given horizontal measure of distance from country A to B, or from B to A, with regard to the study of a particular phenomenon. The same is also possible with vertical measures of distance, however, given their qualitative properties it is also possible to study the differing effects from a positive distance, or negative distance. Not negative in the sense that the distance is a negative number, but negative in the sense that the distance represents that the comparison is being made to a country with a worse of situation.

2.6 – Conceptualization of Institutions

From Ghemawat (2001) and Berry, Guillen and Zhou (2010) we can say that Administrative Distance will be a component of an overarching Institutional Distance. This, because we are using the broad view of Institutions. This makes it important that we understand not only administrations, but institutions as well, for the sound development of frameworks utilizing these concepts.

The definition of what institutions are has evolved notoriously over the years. Being used in several areas and with several schools of thought behind it, as well as about a century in on again, off again, prominence, the definitions are certainly varied (Scott, 2014). To open the discussion of their definition, before going further into a historical contextualization, we can use North's (1991) one, falling into Neoinstitutionalism, that Berry, Guillén and Zhou (2010) and Ghemawat (2001) seem to follow, in which Institutions are, in the most general sense, and

in a straightforward manner: What mediates, explicitly or implicitly, the rules of the game (North, 1991).

The very first notable Institutional arguments related to economics originated in Austria and Germany, around the beginning of the twentieth century, surrounding discussions about the scientific method in the social sciences. Gustav Schmoller led this initial movement in Europe, while in the United States, particularly scholars that had had contact with this school of thought, would import it and develop it further. This movement did not only cross national borders, but areas of science, travelling towards Political Science, Sociology (Scott, 2014), and evidently by the existence of such a thing as Institutional Distance, to Management and International Business.

Gustav Schmoller and this early Institutionalism came to the fore essentially to address the simplifications utilized in classical economics, considering that the social setting, as well as historical and cultural aspects, had a part to play in the economic process. Perhaps the biggest critique that they made was to the rational economic agent that made decisions in a calculated and abstract manner. Opposed to Schmoller was Carl Menger, defending classical economic simplifications that could work in abstraction and that were true regardless of time and place. Essentially this early stage of Institutionalism is aiming at rethinking if economic systems are determined or not, if preferences are endogenously determined or exogenously, if the best way to proceed is through simplifying assumptions or through behavioral realism, and if the insights achieved are time invariant or not. The movement develops, and both sides of the argument fail to convince one another of a definitive answer, and classical economics winds up taking dominance in the academic field, up until there is a resurface of Institutionalism in the 1970's, this time in company of Neoclassical economics and in the form of Neoinstitutionalism (Scott, 2014).

Max Weber, at a stage where the debate was more consolidated, winds being who better reconciles both sides, as well as other strands of Institutionalism, despite not referring Institutions in an explicit manner (Nau, 2005). Through Nau's (2005) work we can get a distilled version of Weber's outlook:

“We can identify three main building blocks for this economic theory of institutions:

1. a theory of property rights that describes the individual and the group incentives in the system;
2. a theory of state, since it is the state that specifies and enforces property rights;
3. a theory of ideology that explains how different perceptions of reality affect the reaction of individuals to the changing objective situation” (Nau, 2005).

Other strands, movements and contributions existed meanwhile (Scott, 2014), however, for brevity purposes, after looking at the origins we will look at the resurface that materialized in Neoinstitutionalism. Neoinstitutionalism gave much more attention towards property rights, transaction cost, commons, innovation and evolutionary theory, among issues of the same nature, and it developed at a time when economic theory had a firmer ground. Neoclassical Economics has at this time proven concepts that work without much outside context, using simplifying assumptions and tight restrictions over the conditions of use of particular models, nevertheless falling short in certain aspects and areas in terms of explanatory capacity. In Neoinstitutionalism, many of these “gaps” are attempted to be filled, giving higher weight to the fact that coordination of the economy is not done through markets alone, markets themselves that depend on an existing institutional arrangement, and focusing on the impact of certain economic processes, as well as broadening the conceptualization of the economic agent (North, 1991; Kahneman, 2012; Scott, 2014).

Williamson (1975) furthers and revitalizes the work done by Robert Coase (1937), that albeit heavily cited had generally not been developed further. Coase (1937) had stated that there had to be a cost in the process of using the price mechanism, otherwise firms would never internalize anything. Williamson (1975) furthers that transaction costs were to increase when “the limited human rationality is confronted by heightened complexity and uncertainty”. While Williamson’s (1975) work focusses mainly on the impact of institutions at the firm, or micro level, dealing with how to best match structure to strategy, as well as how to minimize said transaction costs, North (1991) shift his perspective towards a higher level of analysis.

North (1991) defines Institutions as:

“...the humanly devised constraints that structure political, economic and social interaction. They consist of both informal constraints (sanctions, taboos, customs, traditions, and codes of conduct), and formal rules (constitutions, laws, property rights)”.

Transaction costs are still a topic of study for North (1991), but he treats it as a dependent variable of an overarching institutional framework, rather than as an independent variable that would be the predictor of how different forms of governance, or different strategies, are chosen. Thus, transaction costs to the extent that they play a part in shaping any company's behavior, are in themselves determined by this larger institutional framework.

Scott's (2014) definition, first published in 1995, further integrates the several strands present in Neoinstitutionalism, while being seemingly the more contemporaneous of all of them. This not only by its merit, but also, in part, because the book has been continuously updated.

Scott (2014) proposes that the definition of Institution lay on three pillars: Regulative, Normative and Cultural-Cognitive.

- The Regulative Pillar encompasses the more explicit part of institutions, such as “rule setting, monitoring, and sanctioning activities”.
- The Normative Setting encompasses “normative rules that introduce a prescriptive, evaluative, and obligatory dimension into social life”, thus encompassing values and norms.
- The Cultural-Cognitive dimension is the one which Scott (2014) considers the fresher addition his work proposes, having more prominence when looking through the lenses of Organizational Studies or Sociology. This pillar highlights “the shared conceptions that constitute the nature of social reality and create the frames through which meaning is made” being that the hyphenation highlights “the internal interpretation of each individual, that is in turn shaped by external cultural phenomena”.

One of the core ideas surrounding this construction of institutionalism through pillars is that in a healthy institutional setting each of the dimensions works together in harmony. However, in times of institutional destabilization, regardless of its nature, Institutions as a whole might be only held up by one of the pillars, none, or with pillars “pushing” the institutional landscape

in several directions. Point being that the “pillar” metaphor might be beneficial for the construction of models of understanding surrounding institutions.

Evidently, even in Neoinstitutionalism there are several strands of thought that were not included here, ranging from game theory, rational choice theory, culture or cognitive theory, just in the field of economics (Scott, 2014). However, most International Business scholars seem to draw their definitions regarding institutions either from North (1991) or Scott (2014) (as example: Blanc-Brude *et al.*, 2014; Ang, Benischke and Doh, 2015; Hasan, Ibrahim and Uddin, 2017; Parietti, 2017; Quer, Claver and Rienda, 2017; Adamoglou and Kyrkilis, 2018b; Mingo, Morales and Dau, 2018; Kostova *et al.*, 2019). A good way to close this discussion is through the work of Child (2000) that developed further clarity surrounding the topic of institutions. Essentially, the author points out that the discussion around Neoclassical economics or Institutionalism is really one of a low context perspective, versus one of high context perspective.

In a lower context setting the organizational, cultural, institutional or similar other differences are of little importance, as the trend line will average them out eventually, thus forecasting a convergence over time to the most efficient “solution”. In a high context setting there is, on the other hand, the presumption that an organization, culture, institution or other will develop and retain different characteristics based on particular preferences, developmental history, and other circumstantial factors (Child, 2000). This is particularly the strand of thought motivating Whitley's (1992) work, regarding comparative analysis of institutions, conceptualized as business systems, that then Berry, Guillén and Zhou (2010) use as one of the theoretical bases for several of their proposed distances, including the Administrative Distance.

High and low context are not necessarily irreconcilable, as there are situations and aspects, which are less, or more, dependent on context. Likewise, there are approaches which benefit of less, or more, contextualization (Child, 2000).

2.7 – The evolution of Distance Measures

When it comes to distance in this context of research, an often cited quote of Zaheer, Schomaker and Nachum (2012) (Ambos and Håkanson, 2014; Duarte and Carvalho, 2018; Kostova *et al.*, 2019; López-Duarte, Vidal-Suárez and González-Díaz, 2019) highlights its relevant nature stating: “Essentially, international management is management of distance.”

The authors then develop their assertion saying:

“Coordinating the activities of units or firms that are dispersed across countries requires dealing with time zones, miles and communication by phone and email. Beyond this, international firms must also face the myriad country differences that contribute to other forms of distance, such as cultural, administrative and economic distance (Ghemawat, 2001).”

This is not without some measure of discordance, as is the case of Maseland (2018), which uses the quote to question distance and propose alternatively institutional overlap⁵, although its usage for its critique seems scarcer.

The first big and influential works related to distance in International Management come from the University of Uppsala. The central article of this body of work is that of Johanson and Vahlne (1977) dealing with the increase in commitments to external markets as there is an increase in knowledge about them, chiefly achieved through practical learning in interacting with said markets, and related to the psychic distance that was present at in the beginning of said interaction. Psychic distance is defined by the authors here as “the sum of factors preventing the flow of information from and to the market. Examples are differences in language, education, business practices, culture, and industrial development.” To note, psychic distance is conceptually tied to the individual, even if accounting for external factors, although its definition varies substantially (O’Grady and Lane, 1996). This said the authors are not particularly interested in the individual decision maker, but rather in the decision-making system. Evidently, previous articles exist that lead up to the one that brings them all together

⁵ Boiled down, slight differences do not matter in this framework, rather, it takes a more binary approach in measuring if two things are the same, or not. The core argument is that similar is just as bad as dissimilar, and that only equal works in fostering international business and making it easier. Two examples might be language or the legal system. It also does away with transitivity.

and develops the model, such as that from Johanson and Wiedersheim-Pau (1975), as well as a family of relevant articles that further developed this framework.

It may be important to note that the Johanson and Vahlne (1977) use of psychic distance is not original. The introduction of the broad concept of psychic distance seems to have been made by Beckerman (1956), as stated by Dinner, Kushwaha and Steenkamp (2019) and in general agreement with several other authors (e.g. Kuo and Fang, 2009; Ambos and Håkanson, 2014; Hutzschenreuter, Kleindienst and Lange, 2016; Vaccarini *et al.*, 2017). In it, Beckerman (1956) states, when regarding specifically psychic distance:

“As well as speculating on the future relative rates of growth in terms of nearness, it would be interesting to speculate on the effects of changes in modes of transport and transport costs. Apart from the general reduction in 'economic distances which would result from a reduction in the costs of air freight, for example, a special problem is posed by the existence of "psychic" distance. It is probable that the manner in which the purchases of raw materials by a firm are distributed geographically will depend partly on the extent to which foreign sources have been personally contacted and cultivated. While the transport costs paid (directly or indirectly) by an Italian entrepreneur on a raw material supplied by Turkey may be no greater (as the material may come by sea) than the same material supplied by Switzerland, he is more likely to have contacts with Swiss suppliers, since Switzerland will be "nearer" to him in a psychic evaluation (fewer language difficulties, and so on), as well as in the economic sense that air travel will absorb less of his time. The growth of air travel and freighting would not only tend to iron out discrepancies arising from sea versus rail transport (as the aircraft travel, in general, along straight lines - or great circles) but would have very interesting repercussions on psychic distances.”

Hofstede's (1980) work is the next big development in distance frameworks. In it, he develops four different dimensions of culture, them being: power distance, uncertainty avoidance, masculinity/femininity and individualism/collectivism. Later, in 1991, he adds long term/short term orientation (Hofstede, 1991 apud Minkov and Hofstede, 2011; Hofstede and Minkov, 2010), and in 2010 he adds indulgence/restraint (Hofstede, Hofstede and Minkov, 2010 apud Minkov and Hofstede, 2011) as a sixth dimension.

Associated with Hofstede (1980) cultural dimensions comes Kogut and Singh's (1988) work. This article chiefly contributes to the operationalization of Hofstede's cultural dimensions through the widely used KS index⁶, which aggregates the several dimensions. In 2012, in fact,

⁶ Kogut and Singh Index

it was still the most common operationalization of Hofstede's cultural dimensions as an aggregate (Zaheer, Schomaker and Nachum, 2012), further given that it works just the same with four dimensions as it does with six by its nature. We have no reason to think that it is not the most widely used index for these purposes in 2019 or 2020⁷. This index has suffered a good measure of critique for several of its limitations, being that Shenkar (2001), which called to attention said limitations, was a high advocate of evolution away from it. However, the implicit behavior of researchers is that the usefulness of the index outweighs said limitations. Most that mention Shenkar (2001), in fact, do so only to highlight that they are aware of said issues present in it (Shenkar, 2012).

To note, Hofstede's model is not the only one to handle cultural dimensions as a way to measure and conceptualize distance. Other models can be found with relevant utilization to merit its discussion here such as Schwartz (1994) or House *et al.* (2004). Other authors utilize survey responses, and not country level data, such as Driscoll and Paliwoda (1997), Luo (2002) and Solberg (2008), although their lesser utilization may be due to the higher work required for their implementation. To the extent they are not further developed here, it is by considerations of space and time.

2.8 – Institutional Distance Models

As Institutional Distance models are not the focal point of this dissertation, we will focus on the two that more prominently feature Administrative Distance. Ghemawat's (2001) CAGE model, and Berry, Guillén and Zhou's (2010) Institutional Distance model.

Ghemawat's (2001) CAGE model is the first widely adopted Institutional Distance model that incorporates the measurement of Administrative Distance, although the author does not specify what type of distance the model as a whole measures in his article. It seems, however, implicit that this is supposed to be "the" distance, and can be easily interpreted as being referent to Institutional Distance. The model can have this position since it is made clear in the

⁷ In 2019 Google Scholar reports 435 citations, out of 7222 in its total lifetime. If we divide the citations by the number of years the article has, we would have only about 225. In 2018 the article saw 445 citations, and in 2017, 438 citations. 2012 yields 388 citations. For 2001, it has 109 citations.

article that this model requires heavy consideration not only about the type of industry that it analyses, but also about the internal characteristics of the company. This is, in addition to careful consideration regarding the four distance it uses, them being: Cultural, Administrative, Geographic and Economic. Similarly, no method for combining the four is given in the article.

However, if we may say that Ghemawat's (2001) CAGE model is the first Institutional Distance model, particularly with respect to one that incorporates Administrative Distance, it is also true that authors use both North's (1991) and Scott's (2014) institutional frameworks to calculate Institutional Distance (Abdi and Aulakh, 2012; Ang, Benischke and Doh, 2015).

Berry, Guillen and Zhou (2010) present the next notable inclusion of Administrative Distance in their Institutional Distance model. While it borrows some elements from Ghemawat (2001), at least in concept, while others are quite novel. It presents, fundamentally, the four CAGE distances, even if measure with different variables, with further contextualization provided by five more distances. A new measurement method is also used in the form of the Mahalanobis Distance (Mahalanobis, Bose and Roy, 1937), which will be presented in more detail later on.

Berry, Guillén and Zhou (2010) base their theoretical approach on “three conceptualizations of cross-national institutions”. These three conceptualizations are:

1. National Business Systems

- a. Origin: Whitley (1992)
- b. Definition: "particular arrangements of hierarchy-market relations becoming institutionalized and relatively successful in particular contexts" (Whitley, 1992, p10). Their business systems can differ in terms of economic, financial or administrative practices.
 - i. These differences originate because of: demographic, geographic, cultural and political institutions.

2. Implications of differences in national systems of governance

- a. Origin: La Porta *et al.* (1998), Henisz and Williamson (1999)
- b. Definition: National Governance Systems refer to the “Set of incentives, safeguards, and dispute-resolution processes used to order the activities of

various corporate stakeholders” (i.e. owners, managers, workers, creditors, customers, society)

- i. These differences originate in administrative and political institutions that make different stakeholders more or less powerful depending on their history.

3. National Innovation Systems

- a. Origin: Nelson and Rosenberg (1993)
- b. Definition: Different configurations of institutions foster the development of technology and innovation.
 - i. Since different countries can produce knowledge at different levels, this relates, limiting or enhancing, their ability to leverage this asset by being connected to other countries.⁸

They do this following the recent institutional theorizing in the field of international business of Jackson and Deeg (2008) and Pajunen (2008). With this in mind, they propose the following distances: Economic, Financial, Political, Administrative, Cultural, Demographic, Knowledge, Connectedness and Geographic.

On measurement, Berry, Guillén and Zhou (2010) highlight five desirable properties of a distance measurement: symmetry, non-negativity, identification, definiteness, and triangle inequality. The authors point to the issues present in the Euclidean Distance or Euclidean Squared Distance, being that:

1. It does not take into consideration the correlation between the variable indicators used to computing it, which when it happens means they are capturing the same characteristics, giving more weight to the same variables;
2. Does not take into consideration the variance of the variables;
3. It is sensitive to the scale of the measurement. (Berry, Guillén and Zhou, 2010)

⁸ The above 3 points are almost word for word taken out of Berry (2010). Quotation marks were not used since the text was rearranged in order to enhance clarity, and slightly changed, although overall retaining its fundamental meaning.

The solution presented is the Mahalanobis distance (Mahalanobis, Bose and Roy, 1937), which we will follow. It meets the five criteria listed previously and surmounts the three problems identified, taking into account the variance-covariance matrix and being scale-invariant.

2.9 – Administrative Distance as a component of Institutional Distance Models and its Conceptualization

We chose to aggregate this chapter, taking the equivalent form for Administrative Distance that 2.6 and 2.8 took for institutions. This because the literature and ground to cover in regards to Administrative Distance is much shorter, despite much more important for our work. Because of this, it seems also beneficial that both are done in tandem, due to the relatively scarce literature when it comes to Administrative Distance, thus theory benefiting from the insights practical application might render by inference.

Regarding Administrations, here in the sense of: That which Administrative Distance measures the distance of, there is not in the same sense as for institutions such a large body of work to draw from. To the best of our knowledge, there is not any large current of Management, International Business, Economics or some other social science that is called something like “Administrationism”, but rather what has existed is the development of “administration” as a concept, and certainly as a practice. Further, a problem that encroaches on us getting a clear conceptual clarity of what is meant by “Administrative”, in “Administrative Distance” is the fact the “Administration” as a standalone concept can refer to many different things. Fayol's (1981) work, for example, despite retaining the same name is not concerned with the same type of subject as ours, much more focused on company and employee specific processes. On the other hand, some parallels can be traced to Scott's (2014) Regulative Distance. In this later case, rule setting institutions, as an example, are quite conceptually connected to the legal framework, which Berry, Guillén and Zhou (2010) use to measure, in turn, Administrative Distance.

Specifically, Berry, Guillen and Zhou (2010) define Administrative Distance as “Differences in bureaucratic patterns due to colonial ties, language, religion, and the legal system” however,

that leaves open some questions as: Bureaucratic patterns in what? Why just due to these variables? In addition, in some cases, why due to some variables at all, such as religion, which does not appear to have a firm supportive body in the referenced literature (Whitley, 1992; Lubatkin *et al.*, 1998; Henisz and Williamson, 1999; Jackson and Deeg, 2008; Pajunen, 2008; Guler and Guillén, 2010).

All of the variables are constant, but language was dropped in the 2017 update provided by the authors due to the lack of reliable data. The following articles are referenced as supporting their conceptualization and measurement of Administrative Distance:

1. Theoretical sources in the institutional literature: Whitley (1992) apud Berry, Guillén and Zhou (2010), Henisz and Williamson (1999), Henisz (2000).
2. Examples of empirical studies in the international business literature: Lubatkin *et al.* (1998), Guler and Guillén (2010).

For Berry, Guillén and Zhou (2010) Administrative Distance is also just as stable through time as Geographical Distance, which seems quite a hard assumption. Countries that were once united, not long ago, or have otherwise had a shared history that ventures beyond religion or colonialism, are contenders to identify administrative aspects that affect Administrative Distance that this version of it might have neglected. In that vein, the commonality of second languages might be relevant, if language (interpreted as the native language) is.

As an example to illustrate the possibility, and reasonableness of this, in principle, one needs only to look at former USSR countries, in which many people (still) speak Russian as a secondary language. In a very good sense, considering whichever concept of “administrative” one wants, and further considering which variables are generally used, the impact of the USSR goes, in principle, and as an example, with certain parts of it largely neglected. That looks even more likely to happen even if we take the concept of administrations as bureaucracy.

In fact, one of the Berry, Guillén and Zhou’s (2010) Administrative Distance sources, Whitley (1992) is precisely about how different historical factors, in different countries relatively close

to one another, and with several similar characteristics⁹, have contributed to quite different institutional, and administrative systems.

Ghemawat (2001) does not directly define Administrative Distance. He goes straight into listing the variables that comprise Administrative Distance and why they are important. They are: Colonial Ties, Shared monetary or political association, Government policies and Institutional weakness.

The following paragraph is what he writes regarding institutional weakness, and from it we might be able to expand:

“Finally, a target country’s weak institutional infrastructure can serve to dampen cross-border economic activity. Companies typically shy away from doing business in countries known for corruption or social conflict. Indeed, some research suggests that these conditions depress trade and investment far more than any explicit administrative policy or restriction. But when a country’s institutional infrastructure is strong—for instance, if it has a well-functioning legal system— it is much more attractive to outsiders.”

Then, if need be, we might be able to replace “institutional weakness” with such variables as corruption, social unrest, well-functioning legal system and property rights.

However, Ghemawat (2001) goes more in depth to predict industries/variables that predict a strong influence of Administrative Distance, them being: producers of staple goods (electricity, water), producers of other “entitlements” (drugs), large employers, large suppliers to governments, national champions, vital to national security, exploiters of natural resources, subject to high sunk costs.

By the situation that the CAGE model (Ghemawat, 2001) has only: Cultural, Administrative, Geographic and Economic Distances, we can infer to some extent that the political dimension present in Berry, Guillén and Zhou (2010) is integrated within the CAGE’s Administrative Distance. For Ghemawat (2001) language remains a variable of interest, as well as religion, but the author considers them to be part of Cultural Distance. In Ghemawat and Hout (2017) the

⁹ Japan, Taiwan and South Korea, specifically.

authors make reference towards the legal system being a relevant variable of Scott's (2014) Regulative Pillar, however never explicitly addressing it towards Administrative Distance.

3 – Study Proposal for the Improvement of Administrative Distance

Calling back Shenkar's (2001) illusion of causality, as this works pertains to the development of Administrative Distance, and in it a seemingly lack of substance despite relatively high levels of statistical significance being frequently found (Kuo and Fang, 2009; Blanc-Brude *et al.*, 2014; Quer, Claver and Rienda, 2017; Liu *et al.*, 2019), the issue of the causal mechanism of Administrative Distance is of particular relevance, as it influences the interpretation that can be made out of the results obtained. Additionally, if there is an omission of relevant variables, then their inclusion might either cause a drop in significance of the variables we had before, as the effects the previously had are now captured elsewhere, or an increase, as the effects they had are now better contextualized, or disentangled from the previously omitted relevant variables (Wooldridge, 2013).

While this later effect seems at first sight less common, to get an empiric sense of it we might only imagine the likely outcome of trying to measure a sportsman ability to score by his or her salary. We might even find that as the salary increases the ability to score, as measured by salary, does not significantly increase. If this were to happen then, we would immediately think something wrong with our model and realize that we were not considering the league at which said sportsman was playing, or his or her position in the game. In this case, then, the inclusion of these omitted variable could possibly increase the statistical significance of salary as a predictor of scoring ability (Wooldridge, 2013, p. 90).¹⁰ It is also just as possible that this illusion may play a role in less than ideal proxy variables. Even in ideal proxy variables, that little alter effectively the results obtained, it is worth the consideration over the nature which interpretation might take if we are not cognizant of this aspect present in a model¹¹ (Wooldridge, 2013, p. 309)

¹⁰ To clarify, the example given is not based on any study, nor are we saying it is the case. It exists simply to improve the conceptual clarity of a possible occurrence at the inclusion of a relevant variable, previously omitted.

¹¹ (i.e.) Interpreting the proxy as opposed to the underlying variable and phenomena.

Then, in the interest of developing the construct of Administrative Distance, and considering its general aim at measuring bureaucratic differences, we propose that Berry, Guillen and Zhou's (2010) variables lack a strong measurement of more tacit bureaucratic, administrative differences, focusing on the other hand on overarching, historical ones. Our choice with Berry, Guillen and Zhou's (2010) model is derived from this being the model that, being widely accepted, seems to be the most encompassing regarding distance dimensions and in this sense permits a higher isolation of Administrative Distance. While Ghemawat's (2001) CAGE model could also be a valid choice, the model's open-ended technical specifications, which can be beneficial in other contexts, here that comes as a disadvantage. Even regardless of that, Berry, Guillén and Zhou's (2010) higher number of dimensions, that permits a better de-interlacing of what it and isn't part of Administrative Distance, together with its usage of the Mahalanobis Distance, makes this model in our view, the most appropriate.

We chose to test variables that measure different facets of administrative and bureaucratic processes within a country in a more direct fashion, which we found in the *Doing Business Report* from the World Bank (World Bank Group, 2020). Evidently, all of these are reflective of a larger whole that is meant to be captured, they are: Ease of starting a business, dealing with construction permits, getting electricity, registering property, getting credit, protecting minority investors, paying taxes, trading across borders, enforcing contracts and resolving insolvencies. We test these variables as an addition to Berry, Guillén and Zhou's (2010) model, rather than a substitution.

Our concerns with the choice was that the proposed variables showed:

1. Prior fit and relation to the study of FDI;
2. Conceptual connection to Administration or Bureaucracies, that is, that are fit to be measured as Administrative Distance;
3. Fit for the lacking aspects of Berry, Guillén and Zhou's (2010) Administrative Distance

While points two and three are expanded on previously, the lack of fit for the lacking aspects of Berry, Guillén and Zhou's (2010) Administrative Distance are the reason we did not choose

broader possible measures of Administrative Distance, such as WGI. The *Doing Business Report* variables seem to better pinpoint the lacking practical bureaucratic aspect we were looking for, and found lacking in Berry, Guillén and Zhou's (2010) Administrative Distance. Regarding general fit for the study of FDI, and particularly inward FDI, the *Doing Business Report* does not lack examples of previous application (Pinheiro-Alves and Zambujal-Oliveira, 2012; Corcoran and Gillanders, 2014; Shahadan, Sarmidi and Jan Faizi, 2014; Hossain *et al.*, 2018; Haliti *et al.*, 2020), although its use within a broader model, as well as distance calculation, as opposed to using the variables in level, seems relatively less common.

This said the issue of aggregation or disaggregation arises, specifically if the *Doing Business Report* variables should be calculated in conjunction with, or separated from, Berry, Guillén and Zhou's (2010) Administrative Distance variables. The ideal would probably be to create an aggregated, holistic variable measuring Administrative Distance. However, due to the fact that we have only access to the distance data from Berry, Guillen and Zhou (2010), rather than the original values, as well as not having access to some methodological decisions that must have been made previous to the Mahalanobis Distance calculation, that option is limited by these constraints¹².

¹² We can take the % of common religion as an example.

The CIA World Fact Book presents at best an incomplete and inconsistent view, even if it is the best that can be done. Its objective seems far from what we are here trying to use it. Therefore, it often presents its variables in differing forms of aggregation. While one country might be indicated to be x% Shia, the other might be indicated to be y% Muslim, despite if we make a quick search realizing that it is predominantly, for example, Sunny.

Secondly, the World CIA Fact Book is itself, in this regard, a compilation of different studies made by different countries regarding the religious identifications of its citizens, which points to different methodologies in use. Thirdly, there is no specification over what constitutes the "same" religion. Are Shia and Sunny the same? Likewise for Catholicism, or Orthodox, since they belong to the same overarching religion, be it Muslim or Christian. In that case, is that a reasonable hypothesis, in which case Europe as a whole will share, mostly, the same religion, and differences for which many wars have been, and are being, fought are neglected as insignificant. On the other hand, they very well might be, if these differences do not play a part when it comes to bureaucratic processes formulation, which we want to access.

In the same vein, what are we to do with Atheists, Other, or Non-Specified. Frequently there is no Atheist component, where with a high degree of confidence, at the very least a residual Atheist following would be expected, which could lead us to see a bias in Other, or Non-Specified. Further, as we are primarily trying to measure impact in Administrative dimensions, through religion, should we then only consider the major religion(s) in any given country, forgoing small differences, particularly in light of inconsistent data.

The *Doing Business Report* Administrative Distance variables in consideration, however, fail to make horizontal distinctions (Ghemawat and Hout, 2017), in the sense that two countries might have the same hardship score in regards to opening a business (e.g. measured in days, e.g. 10), however through very different processes, which will cause hardship for a newcomer that has to adapt to them. So while theoretically founded that they should capture relevant administrative results, it is not evident that they will nullify the explanatory power of the previously used variables, if they can indeed offer a satisfactory statistical significance. Instead, the *Doing Business Report* variable offer a much more vertical, qualitative approach. For that reason we will from now on simply call it Vertical Administrative Distance. In this context, we will also call Berry, Guillén and Zhou's (2010) Administrative Distance simply, Horizontal Administrative Distance.

To proceed with this study we decided to take a comparative approach, using a baseline study in which Administrative Distance showed itself to be of high significance. From that baseline study, we will make the restrictions necessary for the introduction of the Vertical Administrative Distance Variable. This is, then, the central part of our work and for which we will delineate our hypothesis.

Due to our limitations, but in the interest of a preliminary analysis, after that, we will proceed with a brief study of the inner components of the Vertical Administrative Distance, and of an Integrated Administrative Distance.

This is to say, while valid decisions can be made regarding the measurement of this variable, component of Administrative Distance, the decisions taken are not straightforward, as they would much more be in evaluating the distance in two GDP's per capita.

Berry, Guillen and Zhou (2010) have neglected to address any of these considerations, anywhere we could find. This, in itself, is reasonable if the larger objective is to put out a working distance framework, and the non-inclusion of these issues certainly sidesteps much discussion that could be surrounding it, for the tens of variables that their nine distances require. However, and as the model had a good amount of success and adoption, we propose that further study in this area is required, so that the treatment of the component variables has some measure of consistency and is able to be replicated. This as well, improves the interpretability that studies can have, by virtue of having them stand on firmer ground.

We have attempted to emulate the calculation that might have been made in Berry, Guillén and Zhou's (2010) Administrative Distance, but have failed in that regard to reach anything over about 0.2 correlation, under differing assumptions, in the line of what we discussed in this footnote.

3.1 – Hypothesis Formulation

We believe all of the *Doing Business Report* variables are consistent with the literature, being that they also may benefit on two fronts. The first being that they, or better, their component variables, are more easily observable, in the sense that most can be measured in days or monetary value. The second is that they more likely zero in on the specific bureaucratic processes we are looking for, which are the ones associated with business. This, together with the conceptual lacking of tacit bureaucratic differences within Berry, Guillén and Zhou's (2010) model, we believe there is solid ground for the Vertical Administrative Distance variable to prove significant.

Further, high correlation between Administrative Distance as measured through the *Doing Business Report* variables and Berry's variables, despite expected at first sight is not certain or even necessary. This for two key reasons: firstly, as we have stated, Berry, Guillén and Zhou's (2010) variables seem closer to measuring horizontal differences, meaning, type of legal system, or of religion, does not say anything about the effectiveness that said systems have, but only about them being different systems. The *Doing Business Report* variables function in the opposite sense, giving no heed to how a bureaucratic system works, but looking mostly at the effectiveness of its results. Secondly, high correlation despite expected is not necessary since both of these are highly reflective variables, given that the nature of administrative or bureaucratic processes is not one directly observable. A low correlation might indicate simply that we are achieving a reflection from quite a different angle to the one we had before, being that we can be in fact observing the same thing.

Likely, the worse scenario for the fit of *Doing Business Report* variables is one of *too high* correlation, meaning that they in fact capture not much else, with the *caveat* of course that they should have a statistical significant outcome, preferably improving the model, whether disaggregated, or aggregated, if possible, which becomes further unlikely with high correlations. This, as well, has the possibility to split the significance of Administrative Distance, if both variables are separated, without adding much else (Wooldridge, 2013).

In line with the analysis of components of the construct of Administrative Distance itself, the question also arises, and is one of the critiques of Shenkar (2001), as well as having been approached by Beugelsdijk, Ambos and Nell (2018), over the reasonableness of aggregating these different variables to create a dimension. The process of utilizing the Mahalanobis Distance solves many of our issues, more specifically that of correlation and scale, but we are nonetheless joining, and giving equal weight when we do so, to the different lower level variables we are aggregating.

Because of the lack of access to Berry, Guillen and Zhou's (2010) primary variables, we will choose to hold Vertical Distance as separated, except in the later case of a preliminary investigation onto an Integrated Administrative Distance. As we cannot test our new model against nothing, we will use FDI as the independent variable. Therefore:

H1: Vertical Administrative Distance, as measured by the *Doing Business Report* variables, is statistically significant in explaining FDI, having a negative relation.

H2: Vertical Administrative Distance, as measured by the *Doing Business Report* variables, has low correlation with Administrative Distance, as measured by Berry, Guillen and Zhou's (2010).

Afterwards, we supplement these hypothesis by further testing the inner Vertical Distance variables, as well as investigating a possible placeholder solution for unification of these two Administrative Distances, until it is possible to proceed to an independent calculation of Berry, Guillén and Zhou's (2010) component variables.

The final objective would be the improvement of the model overall. That, however, is pending at least partial confirmation of H1, showing evidence that indeed, the variable belongs in the model, where H2 is concerned in its turn with the variables nature within the model. However, the "general improvement of the model" is an arguable hypothesis, in the sense of not providing a clearly cut indicator by which we can test it. We chose to include it nevertheless, and its evaluation, while with certain subjective characteristics, will be made by a holistic evaluation of all models tested, and the effects that the addition of Vertical Administrative Distance have onto them, either as a separated variable or as part of an Integrated Administrative Distance.

4 – Methodology

As our study regards the distance model itself, specifically the development of Administrative Distance, iterated primarily from Berry, Guillen and Zhou (2010), we need an independent variable to test it against, which as we have said, we have chosen to be FDI. Because of that we have chosen to use a baseline study, Duarte and Carvalho (2018), in which Horizontal Administrative Distance showed a particularly high significance.

We will start with presenting the independent variable, as well as the dependent ones. This section will be less in depth than Duarte and Carvalho's (2018) study, since except for the Vertical Administrative Distance variable, the rest are the same, and our key interest is in said variable. We have chosen to call them by their usual names, but in a sense, everything *but* the two Administrative Distances we will be studying are controls. Next, we will treat the replication of said study, while after we will conduct ours, which required, because of the availability of consistent *Doing Business Report* data, both in time and subject, a reduction of scope. This reduction of scope left the regression with less precision, but the same overall characteristics regardless. Fortunately, as well, we note that for any study focused on more recent years, there would not be this problem in terms of availability of subjects (countries) and time in the *Doing Business Report*, as to propose a new difficult to access data source would prove of less usefulness.

4.1 – Dependent Variable

Our dependent variable is the Foreign Direct Investment Stock in Portugal. While corporate level data would be best, for the study of Foreign Direct Investment, permitting both an aggregated scope in the study, as well as a disaggregated one, due to difficulty of access to said data, as well as privacy concerns worsening this access, Duarte and Carvalho (2018) utilize country level data. Specifically, the Bidirectional FDI data present in UNCTAD (as well as Banco de Portugal, and OECD). Both decisions to measure stock, and not flow, as well as the utilization of UCTAD as a database, are defended more extensively in the original work (UNCTAD, 2017; Duarte and Carvalho, 2018).

The values present in UNCTAD Stock measure, specifically: the percentage of capital that foreign multinational enterprises (MNE's) have in Portuguese companies, as well as reserves including retained profits, plus the net debt these Portuguese subsidiaries owe to said foreign MNE. (UNCTAD, 2017)

4.2 – Independent Variables

The independent variables are the 9 dimensions proposed by Berry, Guillen and Zhou (2010) – administrative, cultural, demographic, economic, financial, geographical, connectedness, knowledge and political distance – as well as the added controls of Exchange rate and Language, and finally the Vertical Administrative Distance variable as conceptualized by the *Doing Business Report*, while dropping language as a control.

All these distances have been calculated using the Mahalanobis Distance, the output that Berry, Guillen and Zhou (2010) provide us, calculated taking into account the observations present in different times, being that we followed the same procedure for our Vertical Administrative Distance variable. This, because the Mahalanobis Distance may be calculated taking into account the observations of several years, or merely the observation of a single year. The method by which it is calculated, then, can disentangle these distances of the correlations between the variables in one specific period of time, or taking into consideration the scope of time which we are studying (Mahalanobis, Bose and Roy, 1937; De Maesschalck, Jouan-Rimbaud and Massart, 2000; Berry, Guillén and Zhou, 2010). The only exception to the application of the Mahalanobis Distance method is the Geographic Distance, which uses the “great circle distance between the geographic center of countries” (Berry, Guillén and Zhou, 2010).

Regarding the Mahalanobis Distance, the calculation it employs is the following (Mahalanobis, Bose and Roy, 1937; De Maesschalck, Jouan-Rimbaud and Massart, 2000; Berry, Guillén and Zhou, 2010):

$$d(x, y)^2 = (x - y)' A^{-1} (x - y)$$

- \mathbf{x} and \mathbf{y} are the vectors of the variables of which we are calculating the distance, present also in its transposed form;
- A^{-1} is the inverse of the covariance matrix of M such that:
 - $M = n \times p$ such that:
 - n represents the lines of each country in each year;
 - p represents the columns of for each of the variables.

On a more technical note regarding the Mahalanobis Distance, this calculation can be done in excel, although there might be more sophisticated software that allows for its automated calculation. In excel the calculation is not done holistically, but in steps, with $(x-y)$ and A^{-1} first, then the production of $(x-y)^T$, then $(x - y) A^{-1}$, and finally $((x - y) A^{-1})(x - y)^T$. While most of these steps only require simple commands, a couple of functions, and vector based calculation by the program, the later if not automated will require manual adjustment to each of the country/year pairings. Excel then, in this manner, becomes impractical for larger datasets, with several hours being lost in this manual adjustment per calculation. This occurs because of the need of simultaneous horizontal and vertical movement of the formulas, in the calculation. To not have this problem, we made a very crude and simple macro that does this. We will leave it as a footnote for anyone interested in calculating the Mahalanobis Distance in this manner, in order to facilitate its implementation in the future¹³. It needs only to be copy and pasted within an excel module, present within the sheet and workbook, its values adjusted and ran.

¹³Sub Mahalanobis()
 Dim B, RowB, RowE, RowB1, RowE1 As Integer
 'These represent comment and are not part of the program
 'Beginning and end table
 RowB = 3
 ColumnB = 83
 'Beginning and end of transposed table
 RowB1 = 372
 ColumnB1 = 2
 'For cycle creates the Mahalanobis Output. It will be located in column 102, and between rows 3 and 350
 For B = 3 To 350
 'Extend ... to the number of variables present in the calculation
 Cells(B, 102) = Cells(RowB, ColumnB).Value * Cells(RowB1, ColumnB1).Value + Cells(RowB, ColumnB + 1).Value * Cells(RowB1 + 1, ColumnB1).Value + ...
 RowB = RowB + 1
 ColumnB1 = ColumnB1 + 1

Since our baseline study had, on occasion, linearly interpolation variables to handle occasional missing values, this is of course present in our replication, since we wanted to achieve the same database. However, it also also occurs occasionally in the Vertical Administrative Distance variable. Usually, it was realized in a situation where one variable was missing one or two years, on an otherwise complete set of observations, or there was the odd missing value in the middle of the series. In the second case, we also realized linear interpolation, on the first case, we realized what can be called an inverted linear forecast to estimate the value of previous entries. This can unduly reduce variance in the variables where it was realized (Wooldridge, 2013), however due to the low amount of variables that suffered this treatment we believe the distortion is not likely to be of too high of a significance. The reasoning was simply that, *in extremis*, excluding a country that might have 53 observations for the construction of the *Doing Business Report* Administrative Distance was not reasonable for the lack of the 54th¹⁴.

We will present now a brief overview of the Horizontal Administrative Distance variable, as well as presenting its sources, as updated in 2017, as well brief look onto the control variables and a slightly more comprehensive view of the *Doing Business Report* Administrative Distance variable. Further information regarding the rest of Berry, Guillén and Zhou's (2010) variables may be found in their article.

Berry, Guillén and Zhou's (2010) Administrative Distance is measured through:

- Colonizer-colonized link: whether dyad shares a colonial tie
 - From: CIA Factbook
- Common religion: % population that share the same religion in the dyad
 - From: CIA Factbook
- Legal system: Whether dyad shares the same legal system
 - From: La Porta *et al.* (1998)

The two explicit control variables used for our baseline study, are: Common Language and Exchange Rate. The first is included since Berry, Guillen and Zhou (2010) removed it from their

Next B
End Sub

¹⁴ The *Doing Business Report* Administrative Distance variable is built using 9 variables over 6 years.

model, for lack of reliable access, however for the Portuguese language and in the countries had in our sample there are no questions as to what countries primarily speak Portuguese or not. The second is included due to evidence found by Buckley *et al.* (2007), having been obtained from the International Financial Statistics, pertaining to the International Monetary Fund, regarding a likely impact of the Exchange Rate in FDI¹⁵.

The data regarding the Vertical Administrative Distance, as conceptualized by the *Doing Business Report* database was obtained through the Doing Business Database under “Historical Data Sets- Custom Query”. The variables included are: ease of starting a business, dealing with construction permits, registering property, getting credit, protecting minority investors, paying taxes, trading across borders, enforcing contracts and resolving insolvencies, as we have said, under the 06-15 methodology. The ease of doing business score was not considered for this study since the methodology changes are more frequent and its earliest value starts in 2010. Further, we achieve the benefits with the Mahalanobis Distance regarding covariance control, and scale control, the formula with which it is calculated being expressed previously. These variables are themselves calculated from other component variables, which we will detail below, as well as use for a further look into the effects present within the *Doing Business Report* Administrative Distance:

- Ease of starting a business:
 - Time, in days, separated for men and women;
 - Cost, in % of income per capita, separated for men and women;
 - Procedures, in number, separated for men and women;
 - Paid-in Minimum Capital, in % of income per capita.
- Dealing with construction permits:
 - Time, in days;
 - Procedures, in number;
 - Cost, in % of warehouse value.
- Registering property:

¹⁵ To note, one of Berry, Guillen and Zhou's (2010) appeal, certainly among others, is one of scope, thus being that the model certainly has scope to encompass, perhaps, the exchange rate, if it proves to be an effective control. This, because, at first sight, it might very well be a welcome addition to perhaps the economic distance, or eventually the financial distance.

This hypothesis denied as plausible, which might very well be, still a consideration over a mechanism of friction regarding the short term Exchange Rate fluctuations, among other similar hindrances to business, even still within Exchange Rate issues, such as over or undervaluation, seem at first sight worth considering.

- Time, in days;
- Procedures, in number;
- Cost, in % of property value.
- Getting credit:
 - Strength of legal rights, in 0-10 index;
 - Depth of credit information, in 0-6 index;
 - Credit registry coverage, in % of adults;
 - Credit bureau coverage, in % of adults.
- Protecting minority investors:
 - Extent of disclosure, in 0-10 index;
 - Extent of director liability, in 0-10 index;
 - Ease of shareholder suits, in 0-10 index;
 - Strength of investor protection, in 0-10 index.
- Paying taxes:
 - Payments, in number per year;
 - Time, in hours per year;
 - Total tax and contribution rate, in % of profit;
 - Profit tax, in % of profit
 - Labor tax and contributions, in % of profit;
 - Other taxes, in % of profit.
- Trading across borders:
 - Documents to export, in number;
 - Documents to import, in number;
 - Cost to export, in US\$ per container (deflated);
 - Cost to import, in US\$ per container (deflated);
 - Time to export, in days;
 - Time to import, in days.
- Enforcing contracts:
 - Procedures, in number;
 - Time, in days;
 - Cost, in % of claim.
- Resolving insolvencies:
 - Outcome, 0 as piecemeal sale and 1 as going concern;
 - Time, in years;
 - Cost, in % of estate;
 - Recovery rate, in cents on the dollar;
 - Strength of insolvency framework, in 0-16 index;
 - Commencement of proceedings, in 0-3 index;
 - Management of debtor's assets, in 0-6 index;
 - Reorganization proceedings, in 0-3 index;
 - Creditor participation, in 0-4 index.

The years of 2003 and 2004 had to be dropped since there was no satisfactory available data for those years, being that we chose to use Duarte and Carvalho's (2018) model 1a, with the needed added restrictions, to have it function as a benchmark. Replication was done, however, for all models. Countries with less than 50% of observations were dropped, those being Brazil, Japan, Malta, Mexico and the United States of America, variables that in aggregate for the remaining countries had less than 90% observations present were also dropped. Only "Getting electricity" qualified for that, that being the reason we have not listed the variable in the section above. In most cases, we had to do the forecast to include 2005 to try to limit the time constraints of our study, and no more. Of all the countries included, only Cyprus stands out as having relatively fewer observations for the period in question. Additionally, the control of language was dropped, as it appeared to not be significant as a control, being that Berry, Guillén and Zhou (2010) dropped it from the Administrative Distance they measured, where it might have been better suited.

Of these variables, those that present perfect collinearity were also removed from our study, this having happened only in the situation of certain variables that had no distinction between men and women, within our sample, despite having different variables for each. As we will now treat the Doing Business Report Administrative Distance with more detail, as there are two layers of aggregation, meaning three levels of variables we can be talking about, to avoid confusion we will set the names we will call them clearly here. If we are talking about the untreated first level variables, we will call them "Vertical Distance component variables", second level (such as: Paying Taxes), we will call them "Vertical Distance pillar variables", being that the third level variable is the *Doing Business Report* Administrative Distance we have been talking about. The aggregation method from second to third level is the Mahalanobis Distance, as we have mentioned, while from first to second level is the aggregation methodology the World Bank has used in building the *Doing Business Report* (World Bank Group, 2020).

Due to the restrictions names above the baseline study ranges from 2003 to 2010, while ours ranges from 2005 to 2010, being that the baseline study has included the following countries: South Africa, Germany, Angola, Saudi Arabia, Australia, Austria, Belgium, Brazil, Canada, Cyprus, Denmark, Spain, United States of America, Finland, France, Netherlands, Ireland,

Iceland, Italy, Japan, Lithuania, Luxembourg, Malta, Morocco, Mexico, Mozambique, Norway, New Zealand, United Kingdom, Czech Republic, South Korea, Sweden, Switzerland, Venezuela.
Our study has the countries except those mentioned above.

5 – Results

5.1 – Baseline Study

As stated, our baseline study uses FDI data collected from UNCTAD’s bilateral trade statistics, being that the dependent variable is the stock of FDI in Portugal, deflated by the Portuguese deflator.

The panel diagnostics tests point to a GLS, Random Effects, regression, in both models, as well this regression specification being consistent with expectation and theory (Wooldridge, 2013). An OLS Pooled regression would suggest no differences among individuals (countries), where a Fixed Effects would not allow the existence of time invariant variables, such as the ones measured through Administrative Distance or Geographic Distance. VIF tests show no significant collinearity. Tables 10 through 17 show the descriptive statistics, correlation matrix, and models 1a, 1b, 2a and 2b. Models 1 have data from 2003 to 2010, while models 2 have data from 2003 to 2015 with some relaxed assumptions. Models “a” have contemporaneous variables, where models “b” have an offset of one year for time-variant variables. In the tables, “d”, when alone, will always stand for “distance” as well as “dist”, “XR” stands for “exchange rate” and “Lang_PT” stands for “Language – Portuguese”. Other abbreviations hold their inferred meaning.

As such, the model takes the form:

$$FDIstock_{it} = \beta_0 + \beta_1 Admin_Dist_{it} + \beta_2 Cultural_Dist_{it} + \beta_3 Demographic_Dist_{it} + \beta_4 Economic_Dist_{it} + \beta_5 Finance_Dist_{it} + \beta_6 Geographic_Dist_{it} + \beta_7 Connect_Dist_{it} + \beta_8 Know_Dist_{it} + \beta_9 Political_Dist_{it} + \beta_{10} Lang_PT_{it} + \beta_{12} XR_{it} + \varepsilon_{it}$$

The control of language was added, as stated, since it was removed from Berry, Guillen and Zhou (2010) model in their 2017 update of their variables, present at Wharton’s University of Pennsylvania website¹⁶, due to the fact that its original source, the CIA Fact book, was deemed to be inconsistent in the way it reported the speakers of languages. The authors point in their

¹⁶ <https://whartonmgmt.wufoo.com/forms/distance-data-downloads/>

update that due to high correlations in the variable components of the measure of distance it was a part of, their Administrative Distance measure showed only a small result due to this change. This control is binary, however that is mitigated by the fact that the countries towards which it is positive, in our sample, unambiguously speak predominantly Portuguese, being that all others unambiguously do not.

Data about these types of variables is hard to find, and easily noisy, particularly if we are searching for data that includes several countries. It generally comes from country level census, which frequently take different approaches measuring it, together with different questionnaires that can influence the values obtained. They are, however, important variables, and as we can see by Tables 12,13, 16 and 17, even data that is likely noisy can hold high statistical significance if it is relevant enough. Language, however, on its own, does not hold any statistical significance. Because of this, it was dropped as a control from the following models, with the added benefit of decreasing the number of independent variables.

We have here gone fairly quickly through the baseline models, which can be found, together with other relevant information regarding them, in tables 1 through 8 in the appendix, being that further information can be found both in Duarte and Carvalho's (2018) article, as well as in Berry, Guillen and Zhou's (2010).

This true baseline construction was meant to ensure that we had a solid foundation with this model before we adapted it for our purposes. Its inclusion in the appendix might also be of interest for a deeper knowledge of how our modification affected it.

5.2 – Vertical Administrative Distance Model

Therefore, with the modifications listed above, we present our model, with diagnostics below, yet still without the Vertical Administrative Distance:

Table 1 - Regression without Vertical Administrative Distance

Model 1: Random-effects (GLS), using 102 observations
 Included 18 cross-sectional units
 Time-series length: minimum 1, maximum 6
 Dependent variable: FDI_Deﬂated
 Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	1101.80	3237.31	0.3403	0.7336	
Admin_Dist	-147.216	53.0967	-2.773	0.0056	***
Cultural_Dist	61.5000	39.3590	1.563	0.1182	
Demographic_Dist	-55.0086	220.476	-0.2495	0.8030	
Economic_Dist	-10.7804	39.5296	-0.2727	0.7851	
Finance_Dist	45.7825	60.4505	0.7574	0.4488	
Geographic_Dist	-0.894076	0.551352	-1.622	0.1049	
Connect_Dist	352.867	495.959	0.7115	0.4768	
Know_Dist	-291.252	172.372	-1.690	0.0911	*
Political_Dist	68.6041	49.4026	1.389	0.1649	
XR	9.38622	4.89375	1.918	0.0551	*

Mean dependent var	5191.547	S.D. dependent var	6718.743
Sum squared resid	2.93e+09	S.E. of regression	5643.798
Log-likelihood	-1020.577	Akaike criterion	2063.154
Schwarz criterion	2092.029	Hannan-Quinn	2074.847
rho	0.079768	Durbin-Watson	1.323472

Source: Author

The model presented above is done for comparative purposes, as it does not have the Vertical Administrative Distance. Panel diagnostics were ran on an Pooled OLS regression which yielded, effectively, P-values of 0 for both the F-Test and Breush-Pagan test, and 0.0925 for the Hausman test, if the *Doing Business Report* Administrative Distance is not included, and 0.2938 if it is, indicating that the Random-Effects Model continues to be the most adequate. The R² of this our model sits at .3644, in comparison to .3653 in the original model, with more years and countries.

As expected, the model just presented takes essentially the form it had before, with t – time, reduced, as i – country, and without Lang_PT:

$$FDI_{stockit} = \beta_0 + \beta_1 Admin_Dist_i + \beta_2 Cultural_Dist_{it} + \beta_3 Demographic_Dist_{it} + \beta_4 Economic_Dist_{it} + \beta_5 Finance_Dist_{it} + \beta_6 Geographic_Dist_{it} + \beta_7 Connect_Dist_{it} + \beta_8 Know_Dist_{it} + \beta_9 Political_Dist_{it} + \beta_{12} XR_{it} + \epsilon_{it}$$

As we can see, in comparison to the originally replicated model, its general characteristics are the same, meaning, same direction of coefficients, generally similar coefficient values; although as expected it has lost quality in the statistical significance of its estimators. Administrative Distance retains 1% statistical significance, the Exchange Rate gets dropped from 5% statistical significance to 10%, Connected Distance loses its 5% statistical significance, and Geographical Distance loses its 10% statistical significance, while Knowledge distance gets 10% Statistical Significance. The only coefficient that changes sign is that of Financial Distance although with a z value of -.35 to one of .75 now, it had very little statistical significance in the original model. Comparing to this model, the introduction of the Vertical Administrative Distance produces the following model, taking the following form, where “Vert_Admin_Dist” stands for the Vertical Administrative Distance and “Hor_Admin_Dist” for its horizontal counterpart:

$$FDIstock_{it} = \beta_0 + \beta_1 Vert_Admin_Dist_{it} + \beta_2 Hor_Admin_Dist_{it} + \beta_3 Cultural_Dist_{it} + \beta_4 Demographic_Dist_{it} + \beta_5 Economic_Dist_{it} + \beta_6 Finance_Dist_{it} + \beta_7 Geographic_Dist_{it} + \beta_8 Connect_Dist_{it} + \beta_9 Know_Dist_{it} + \beta_{10} Political_Dist_{it} + \beta_{11} XR_{it} + \epsilon_{it}$$

Table 2 - Regression with Vertical Administrative Distance

Model 2: Random-effects (GLS), using 102 observations

Included 18 cross-sectional units

Time-series length: minimum 1, maximum 6

Dependent variable: FDI_Deflated

Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	8392.49	3666.16	2.289	0.0221	**
Vert_Admin_Dist	-1241.33	585.729	-2.119	0.0341	**
Hor_Admin_Dist	-138.906	48.8823	-2.842	0.0045	***
Cultural_Dist	42.1893	47.6225	0.8859	0.3757	
Demographic_Dist	-65.5918	187.015	-0.3507	0.7258	
Economic_Dist	12.8057	36.2143	0.3536	0.7236	
Finance_Dist	-10.1907	52.8795	-0.1927	0.8472	
Geographic_Dist	-0.764995	0.466855	-1.639	0.1013	
Connect_Dist	92.1632	486.613	0.1894	0.8498	
Know_Dist	-361.132	152.866	-2.362	0.0182	**
Political_Dist	59.3197	43.8816	1.352	0.1764	
XR	10.2960	4.68432	2.198	0.0280	**

Mean dependent var	5191.547	S.D. dependent var	6718.743
Sum squared resid	2.42e+09	S.E. of regression	5157.078
Log-likelihood	-1010.821	Akaike criterion	2045.641
Schwarz criterion	2077.141	Hannan-Quinn	2058.397
rho	0.047003	Durbin-Watson	1.377382

Source: Author

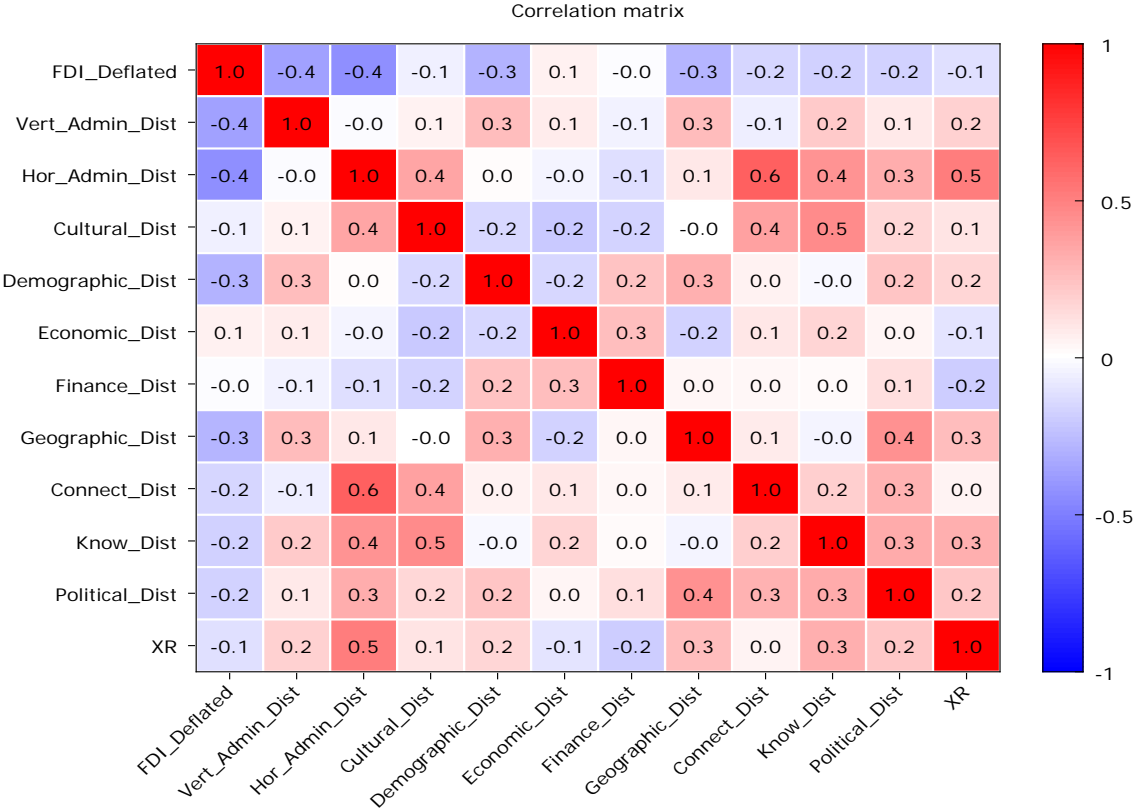
The Vertical Administrative Distance variable, has significance at the 5% level, providing evidence for the confirmation of our first hypothesis (H1). The control of the exchange rate is improved to 5% and the Knowledge Distance has its statistical significance improved to reach the 5% level. Other variables vary when it comes to increasing their statistical significance, some, like Geographical Distance being less affected. In the case of Administrative Distance, it retains its statistical significance of 1%, although it further improves the quality of the estimator, shifting the p-value from .0056 to .0045, which winds up being an improvement of about 30% of an already very significant estimator.

R² increases to .4778 (from .3644), an increase in the explanatory capacity of the model of about 31% with the addition of a theoretically consistent variable. If the F, Breush-Pagan and Hausman tests are now ran in the Random-Effects model, we observe P-Values of 0 again for

the first two tests, and .0276 for the Hausman Test for the first time obtaining a value under .05. We choose to continue the use of the Random-Effects model since several tests before were indicating us to the Random Effects model, as well as, as we have discussed, being the one that is more consistent with the theory underlying this model. A Fixed Effects model would imply that such factors as Geographic Distance have no importance whatsoever, something that is not reasonable to suppose, even if in this current study it does not achieve statistical significance. VIF tests for all models, never show any evidence of possible collinearity.

We can observe that the Horizontal Administrative Distance has a slightly lower coefficient, accompanied by a more significant lower standard error. Specifically, the coefficient (its absolute value) is reduced by close to 6%, while the standard error is reduced by close to 8%. This, in an already very statistically significant estimator leads to said improvement of p-value, which we have stated.

Table 3 - Correlation Matrix with Vertical Administrative Distance



Source: Author

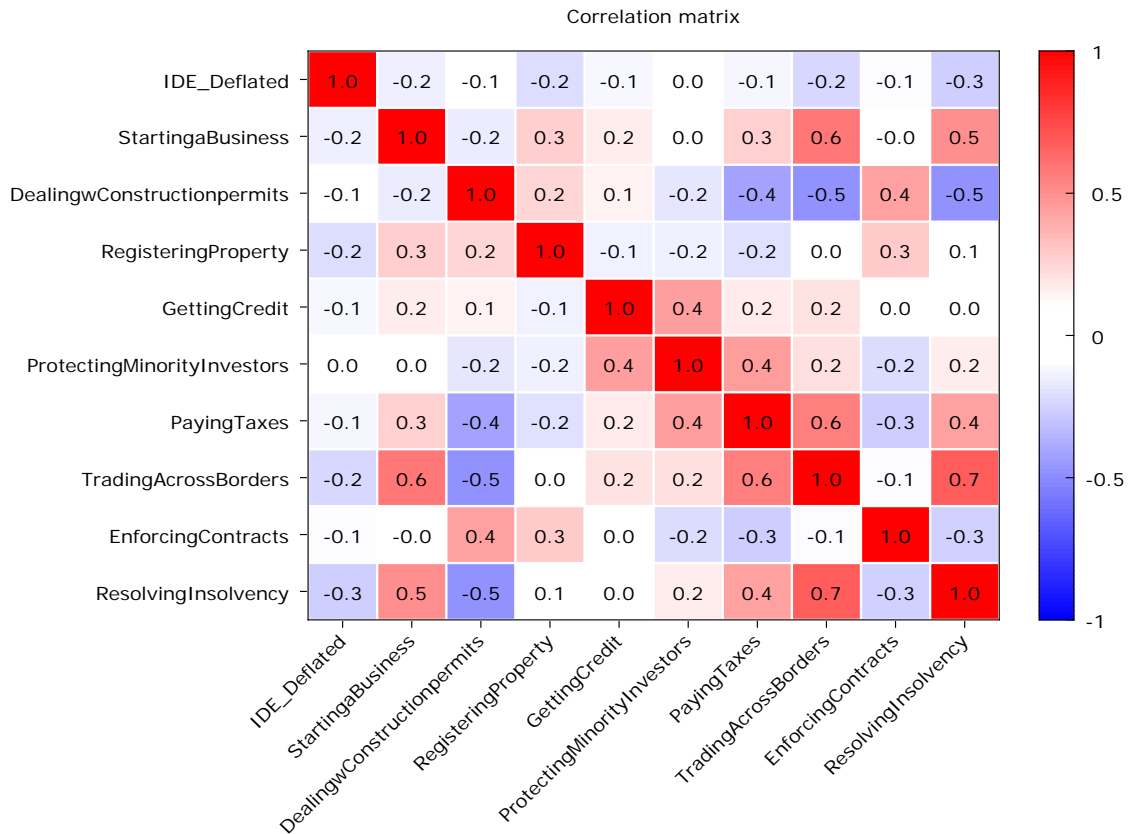
We can see, in table 3, that the correlation between the Horizontal Administrative Distance and Vertical Administrative variable is very low. Since the table only shows one decimal value it appears as zero, although its more precise value is $-.0144$, as to warrant our conclusion that they are essentially independent from each other, thus providing evidence for the confirmation of our second hypothesis (H2). Further, we can see quite high negative correlations between both Administrative Distances and FDI.

Given the listed above, in the effects had on other variables, R^2 , and on the counterpart to Vertical Administrative Distance, of Horizontal Administrative Distance, we can say that there starts to exist preliminary evidence onto the confirmation our third hypothesis (H3). This hypothesis will be further discussed when dealing with the model containing the Integrated Administrative Distance.

These results led us to be interested in studying each component of the Vertical Administrative Distance, running the same regression but with all the nine pillar variables that we are using disaggregated. This so that a higher understanding and interpretation of the model can be made. This proved, however, impossible, since we do not have sufficient observations to run this many variables. However, as a “second best” solution, we decided to run nine different regressions, each with a component of the Vertical Administrative Distance, all in similar random-effects models. The model did not change its characteristics in a meaningful way, so we will present simply the coefficient line of each of these components.

It should be noted, however, that they might be to some extent biased by being separated from the rest. In further studies, if this procedure is to be done, a larger sample size is needed. For this very reason, the correlation matrix accompanies these results close by. This “second best” solution, however, does allow us some primary idea of how the different components interact within the *Doing Business Report* Administrative Distance measure.

Table 4 - Correlation Matrix of FDI and Vertical Administrative Distance Pillar Variables



Source: Author

Table 5 – Isolated Vertical Administrative Distance Pillar Variables

Variable	Coefficient	Std. Error	z	p-value
Starting a Business	-85.047	52.349	-1.625	0.1042
Dealing w/Construction Permits	-193.906	72.617	-2.670	0.0076***
Registering Property	-30.093	81.702	-0.3683	0.7126
Getting Credit	18.562	67.102	1.047	0.2950
Protecting Minority Investors	-276.075	219.093	-1.043	0.2968
Paying Taxes	52.025	78.774	0.6604	0.5090
Trading Across Borders	0.6964	60.120	0.0116	0.9908
Enforcing Contracts	-124.724	139.337	-0.8951	0.3707
Resolving Insolvency	-32.914	39.397	-0.8354	0.4035

Source: Author

Another possible “second best” solution to complement further our view of the pillar variables of the Vertical Administrative Distance is to modify the original model in “general to specific” sense (Campos, Ericsson and Hendry, 2005), which allows for a reduction in the variables of the original model. We did this with the criteria of iteratively remove the variables with the lowest p-value, until all had statistical significance, which then allowed the addition of all the Vertical Administrative Distance Pillar simultaneously. The original reduced model can be seen in table 19 in the appendix.

Table 6 - Reduced Model with Vertical Administrative Distance Pillar Variables

Model 3: Random-effects (GLS), using 126 observations
 Included 23 cross-sectional units
 Time-series length: minimum 1, maximum 6
 Dependent variable: IDE_Deinflated
 Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	0.439722	0.210765	2.086	0.0370	**
Hor_Admin_Dist	-0.497649	0.197846	-2.515	0.0119	**
Cultural_Dist	0.186500	0.117792	1.583	0.1134	
Geographic_Dist	-0.263737	0.168624	-1.564	0.1178	
Connect_Dist	0.0955837	0.0827655	1.155	0.2481	
Know_Dist	-0.245867	0.0832887	-2.952	0.0032	***
XR	0.585996	0.283280	2.069	0.0386	**
StartingaBusiness	-0.0286484	0.0918992	-0.3117	0.7552	
DealingwConstructionpermits	-0.321192	0.147905	-2.172	0.0299	**
RegisteringProperty	0.106761	0.0934340	1.143	0.2532	
GettingCredit	0.0109651	0.104142	0.1053	0.9161	
ProtectingMinorityInvestors	-0.0315223	0.147644	-0.2135	0.8309	
PayingTaxes	0.0815854	0.139953	0.5830	0.5599	
TradingAcrossBorders	0.0233467	0.115839	0.2015	0.8403	
EnforcingContracts	0.0168584	0.106478	0.1583	0.8742	
ResolvingInsolvency	-0.155785	0.102882	-1.514	0.1300	

Mean dependent var	0.166168	S.D. dependent var	0.253272
Sum squared resid	5.216714	S.E. of regression	0.216789
Log-likelihood	21.83184	Akaike criterion	-11.66368
Schwarz criterion	33.71683	Hannan-Quinn	6.772994
rho	0.002145	Durbin-Watson	1.427418

Source: Author

We can see that the correlation with FDI is always negative while internally, several correlations are significant, but not always consistent. This then highlights the importance of

the Mahalanobis Distance calculation, which purges these correlations from the variables above. Dealing with Construction Permits, which on its own does not a particularly high correlation with FDI, stands as the most significant when contextualized through the variables and controls seen above, in both cases. This at most constitutes a preliminary look into the inner workings of the Vertical Administrative Distance, and a higher sample size would be needed in order to try to create a regression with all the pillar variables separated, in a single regression. Tables 20 and 21 show further lower level regression with the component variables of both Dealing with Construction permits, as well as Starting a Business, but to its close to 10% significance in the case where each pillar variable is isolated.

Another issue we realized while studying the Vertical Administrative Distance variable, is the hardness of interpretation due to the Mahalanobis Distance not having a unit, thus producing coefficients that, without closer inspection, can be hard to interpret beyond their general direction, and general magnitude. This led us to consider a unity-based normalization, taking the form of:

$$X_{norm} = (X - X_{min}) / (X_{max} - X_{min})$$

To do this in variables before the application of the Mahalanobis distance simply returns exactly the same distances as output when the Mahalanobis distance method is applied. This is to be expected as one of the functions of the Mahalanobis distance is precisely the normalization of variables. This normalization in the final distance variables, produces in the very same manner an identical regression, albeit with coefficients that are now easier to read, as follows:

Table 7 - Normalized Regression with Vertical Administrative Distance

Model 3: Random-effects (GLS), using 102 observations
 Included 18 cross-sectional units
 Time-series length: minimum 1, maximum 6
 Dependent variable: IDE_Deflated
 Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	0.372674	0.116606	3.196	0.0014	***
Vert_Admin_Dist	-0.229699	0.108385	-2.119	0.0341	**
Hor_Admin_Dist	-0.783887	0.275858	-2.842	0.0045	***
Cultural_Dist	0.0919673	0.103811	0.8859	0.3757	
Demographic_Dist	-0.0760511	0.216837	-0.3507	0.7258	
Economic_Dist	0.0248013	0.0701377	0.3536	0.7236	
Finance_Dist	-0.00636123	0.0330085	-0.1927	0.8472	
Geographic_Dist	-0.591483	0.360966	-1.639	0.1013	
Connect_Dist	0.0641200	0.338547	0.1894	0.8498	
Know_Dist	-0.286093	0.121102	-2.362	0.0182	**
Political_Dist	0.416647	0.308214	1.352	0.1764	
XR	0.717036	0.326226	2.198	0.0280	**

Mean dependent var	0.214157	S.D. dependent var	0.267029
Sum squared resid	3.822877	S.E. of regression	0.204963
Log-likelihood	22.75071	Akaike criterion	-21.50143
Schwarz criterion	9.998247	Hannan-Quinn	-8.746135
rho	0.047003	Durbin-Watson	1.377382

Source: Author

Evidently the coefficients and standard errors are different now, but we might only look at the p-values that are exactly the same, or at the R² which remains exactly the same (0.477639) to see that the model was not altered, as would be expected regardless. This lead us to have a general preference for this normalized model, particularly if we are already handling variables without units, for its added readability at no cost. Any contribution here is of mere form, over substance, but the general improvement in capacity for interpretation seemed worthwhile the addition to this work.

We present now, as well, the isolated pillars of the Vertical Administrative Distance normalized, much in the same way we did before:

Table 8 – Normalized Isolated Vertical Administrative Distance Pillar Variables

Variable	Coefficient	Std. Error	z	p-value
Starting a Business	-0.1865	0.1148	-1.625	0.1042
Dealing w/Construction Permits	-0.3758	0.1407	-2.670	0.0076***
Registering Property	-0.0488	0.1325	-0.3683	0.7126
Getting Credit	0.1151	0.1099	1.047	0.2950
Protecting Minority Investors	-0.4037	0.3870	-1.043	0.2968
Paying Taxes	0.1227	0.1858	0.6604	0.5090
Trading Across Borders	0.0020	0.1710	0.0116	0.9908
Enforcing Contracts	-0.1645	0.1837	-0.8951	0.3707
Resolving Insolvency	-0.1116	0.1336	-0.8354	0.4035

Source: Author

5.3 – Integrated Administrative Distance Preliminary Model

At this junction, despite having handled our primary hypothesis, we wanted to dig deeper into Administrative Distance. Due to the limitations we have stated and faced with Berry, Guillén and Zhou (2010) we were left with a “second best” solution. Because of that, this model then does not take center stage in this work. After giving its presentation, we will then discuss the methodological limitations it faces, as well as its possible benefits.

As we stated before, the Mahalanobis Distance (Mahalanobis, Bose and Roy, 1937) took the following form:

$$d(x, y)^2 = (x - y) A^{-1} (x - y)^T$$

These mentioned limitations forced, as said, a “second best” solution in the way of calculating a Mahalanobis distance (Mahalanobis, Bose and Roy, 1937; Berry, Guillén and Zhou, 2010) of the Mahalanobis Administrative Distances (both Berry, Guillén and Zhou’s (2010) and ours) previously calculated. In the formula used above, the distance between each of the components was present in (x-y), so in order to calculate this “second order” Mahalanobis distance, we have a Mahalanobis Distance input into the equation above such that:

$$d(x, y)^2 = (\text{sqrt}((x - y) A^{-1} (x - y)^T)) A^{-1} (\text{sqrt}((x - y) A^{-1} (x - y)^T))^T$$

This will normalize and remove the correlation between these two distances. As we have already seen a preference for the normalized output and equivalence of these models, when no units in the variables of analysis are involved, we will keep said normalized model.

This calculation implicitly assigns the same weight to both distances, which is not directly evident. Were we to calculate a Mahalanobis Distance based on their respective component variables, we would have a different weight distribution by the fact that Berry, Guillén and Zhou (2010) have 3 component variables, while we have 9 pillars based on several other component variables. However, given that both measures rest on their own theoretical backing, which is conceptually connected but separate, one focusing on historical and overarching measures of Administrative Distance, and the other on practical ones, a simple split of weights has reasonable appeal and logic.

Table 9 - Integrated Administrative Distance Regression

Model 4: Random-effects (GLS), using 102 observations
 Included 18 cross-sectional units
 Time-series length: minimum 1, maximum 6
 Dependent variable: IDE_Deinflated
 Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	0.379311	0.125682	3.018	0.0025	***
Admin_Integrated	-0.406402	0.155449	-2.614	0.0089	***
Cultural_Dist	0.0851763	0.120978	0.7041	0.4814	
Demographic_Dist	-0.0503461	0.205587	-0.2449	0.8065	
Economic_Dist	0.0561117	0.0872796	0.6429	0.5203	
Finance_Dist	-0.0160785	0.0277785	-0.5788	0.5627	
Geographic_Dist	-0.546742	0.356576	-1.533	0.1252	
Connect_Dist	-0.243878	0.334069	-0.7300	0.4654	
Know_Dist	-0.244549	0.116219	-2.104	0.0354	**
Political_Dist	0.213373	0.235905	0.9045	0.3657	
XR	0.402456	0.220141	1.828	0.0675	*

Mean dependent var	0.214157	S.D. dependent var	0.267029
Sum squared resid	4.629264	S.E. of regression	0.224317
Log-likelihood	12.98959	Akaike criterion	-3.979174
Schwarz criterion	24.89553	Hannan-Quinn	7.713177
rho	0.032353	Durbin-Watson	1.409249

Source: Author

Naturally, this aggregation reduces R^2 , but solely to .4004, from .4778 in model 2 and 3, which however is an improvement from 0.3644 in the model 1, with the same number of variables. From another perspective, compared to the model that has the same number of independent variables (that with only Berry, Guillén and Zhou's (2010) variables), we see an improvement in explanatory capacity of close to 10%.

This gives further evidence into the addition of the Vertical Administrative Distance variables providing an improvement over Berry, Guillén and Zhou's (2010) standard model variables, providing further evidence onto our third hypothesis (H3). This Integrated Administrative Distance (Admin_Integrated), consisting on the Mahalanobis Distance of both the outputs of both Administrative Distances addressed here, does have a worse p-value compared to Berry, Guillén and Zhou's (2010) Administrative Distance, despite still retaining its 1% significance, and it relies on values which we cannot replicate, nevertheless being values of a prominent and widely used model. The improvement of the Knowledge Distance, from 10% to 5% statistical significance is retained when the Administrative Distance is aggregated in this manner.

6 – Discussion and Conclusion

As we have stated, the second hypothesis, that of low correlation, being correct was not a given, or even a requirement for a model that would improve on Berry, Guillen and Zhou's (2010) Horizontal Administrative Distance variable (H3) or from a statistically significant Vertical Administrative Distance variable (H1), due to the reflective nature of the construct (Beugelsdijk, Ambos and Nell, 2018). This, because the two types of Administrative Distances seem quite distinct regarding not only the *scope* of Administrative Distances they capture (as they have no correlation), but also the *type* of these Administrative Differences (meaning that the internal characteristics of the variables are different).

That said, H2 *was* confirmed, there being almost no correlation between the two measures of Administrative Distance, R^2 was significantly increased, the Vertical Administrative Distance variable was significant (H1), as well as the control tests having indicated the soundness of the regression itself, being coherent with the existing literature (Wooldridge, 2013).

Both variables are related to Administrative Distance as a concept, coming however from two very different origins regarding its study. Berry, Guillen and Zhou (2010) focus much more on the contextual and historical (Whitley, 1992), while our Vertical Administrative Distance focusses much more on the practical consequences of existing administrative systems in place. As we have seen, they have no correlation with each other, but seem, through the results obtained, to complement and contextualize each other. Not only did the coefficient of Horizontal Administrative Distance become smaller, but also its standard error, which points to some of the effects of the Vertical Administrative Distance as measured through the *Doing Business Report* having been purged from it. Its inclusion in the model seems to de-interlace this capture, creating a more precise estimator of the effects had by this contextual, historical Administrative Distance.

While the Horizontal Administrative Distance does not say much about the difference in *quality*, of administrations, but only *type*, the Vertical Administrative Distance evaluates *only* the quality, saying nothing about type. Companies, when investing in other countries, or performing other activities, seem, in light of these results, to be impacted not only by a

difference in degree of *type* of institutions, but also of *quality*. Berry, Guillén and Zhou's (2010) variables, in Horizontal Administrative Distance, Share of Common Religion, Share of Common Language, Common Legal System (and Common Language if a suitable source is found), do not say anything, explicitly (perhaps only implicitly) about whether a particular administrative apparatus works as well as another. Rather, only that they are likely to be of the same type as influenced by these factors, and that because of this understanding should be easier. Vertical Administrative Distance, on the other hand, is focused solely on results, with complete disregard if said results are obtained in a completely different way.

It is reasonable, then, that companies suffer attrition whether this distance is caused either by different *types* of processes, which they do not know how to navigate effectively, or whether by process of less *quality*, which take longer, and are in general more inefficient, than they are accustomed to and expect.

A first look at a possible Integrated Administrative Distance was made, which gives equal weight to both measures, and relies on a second order Mahalanobis distance calculation. The p-value did decrease, albeit retaining 1% significance, but with the same number of variables we managed a significant increase of the R^2 . Even though the look into an Integrated Administrative Distance was preliminary, this points to the effect not being, in fact, the often dreaded scenario where R^2 is increasing just because we are adding more variables. As our proposed variable had both statistical significance, and was coherent theoretically, this would already be of limited concern, but it is nevertheless uplifting to the proposal of this addition, that nullified the effects of having more variables, we end up, indeed, with a model that better explains the underlying phenomenon we are trying to study, in this case, FDI. This, in turn, gives further strength to the confirmation of H3 and that indeed this addition of Vertical Administrative Distance is beneficial.

6.1 – Contribution

The key contribution we believe we bring to the literature, is the following: if we want to be ambitious with our claim, that the Vertical Administrative Distance, as measured through the *Doing Business Report*, should be included within the overall framework institutional distance

framework, as having both theoretical consistency as well as practical benefit in its application. If we want to be more cautious, in what is probably a more correct formulation, that we have found evidence pointing towards the existence of possible benefits regarding this addition, and further research regarding this issue is warranted, for the benefits, as we have showed, might be quite substantial.

At the very least, we can say with relatively more certainty that inbound Foreign Direct Investment, in Portugal, within our timeframe, is better understood by a supplementation of Berry, Guillen and Zhou's (2010) model with an additional measure of Vertical Administrative Distance, calculated using the *Doing Business Report*, and under this methodology.

The benefits regarding the improvement of the explanatory capacity, either with the model that has Administrative Distance separated or integrated can be seen above, and as we saw they were quite substantial when compared to the previous explanatory capacity of the model overall. Our contribution can also take the form of showing that there is more ground to be covered in regards to Administrative Distance. As we will see next, our study had several limitations, while, as we saw previously, it showed promise of high benefits when it comes to the quality of Institutional Distance models. We hope one contribution can be that of kindling interest in the avenue of research regarding Administrative Distance.

6.2 – Limitations

The specific comparative study we have conducted points to evidence that is narrow: it focusses on one country, on a narrow timeframe, on one subject. For the solidification of the argument that the *Doing Business Report* measure of Vertical Administrative Distance should be included, then, we need studies from different countries, on different timeframes, and different topics, so that we may understand better its behavior.

Additionally, a substantial limitation, particularly when it comes to looking further into the Vertical Administrative Distance Variables was the number of observations we have. Likewise, the inability of replication regarding Berry, Guillen and Zhou's (2010) Administrative Distance variable posed a substantial constraint. For end of this limitation the methodology of

aggregation of the Horizontal Administrative Distance component variables of Berry, Guillen and Zhou (2010) (e.g. share of common religion) has to be known, as it is not self-evident what level of aggregation, or counting with regards to the share, should be used. We attempted to replicate Berry's aggregation methods of component variables for Administrative Distance variables, but were not successful, and could make out even what they might approximately be.

6.3 – Avenues for Future Research

The steep conceptual distinction between these two types of Administrative Distances is more permissive to having them separate. Nevertheless, it would be perhaps ideal that they would be placed as a single variable, as their direction is the same, and they are conceptually connected. We attempted this preliminarily, but further study regarding the matter should be considered, and whether their differences are enough to warrant separation, or not, and then warrant aggregation (Beugelsdijk, Ambos and Nell, 2018).

Further studies should also make use of novel and more extensive data, and again test the improvement of the addition of this Vertical Administrative Distance in conjunction with Berry Guillén and Zhou's (2010) variables.

Regarding Berry, Guillén and Zhou (2010), further investigation should be made in order to handle replicability of results. If this is not the case, issues surrounding the application and inner workings of said models are inaccessible, even before any discussion related to how to measure, transform and identify distance constructs, as well as the implications surrounding that, can be had (Hutzschenreuter, Kleindienst and Lange, 2016). This forbids the consistent addition of missing countries, a quicker yearly update, a fine-grained analysis of the processes, which led to the formulation of said distance values, as well as development of said distance constructs and eventual effective policy recommendation.

Another important issue that became apparent is the lack of consensus regarding what we are talking about, when we speak of distance and distance constructs, something that we already

dedicated considerable time in the beginning of this dissertation, with the objective that any benefit that would be accrued from it might be enjoyed throughout the dissertation.

To end, in a broader view, it seems that much of the way forward in this area of, let's say "dealing with hardships", is highly dependent on a holistic view of it. From clarity of concepts and constructs, to limitation and applicability of metaphors employed, moving onto theoretical model formation and refinement, and moving further towards data selection, treatment, and fit, aggregation, finalizing in interpretation. While certain studies might rely more strongly on some points than others, it seems relevant that none of these areas is completely left out. This is with the objective that further research that intends to use these "hardship to traverse" models instrumentally, with the objective to provide useful interpretations of international phenomena, has higher shoulders to stand on.

7 – Bibliographic References

Abdi, M. and Aulakh, P. S. (2012) 'Do country-level institutional frameworks and interfirm governance arrangements substitute or complement in international business relationships?', *Journal of International Business Studies*, 43(5), pp. 477–497. doi: 10.1057/jibs.2012.11.

Adamoglou, X. and Kyrkilis, D. (2018) 'FDI Entry Strategies as a Function of Distance—The Case of an Emerging Market: Turkey', *Journal of the Knowledge Economy*, 9(4), pp. 1348–1373. doi: 10.1007/s13132-016-0425-1.

Ambos, B. and Håkanson, L. (2014) 'Journal of International Management The Concept of Distance in International Management Research', *Journal of International Management*. Elsevier Inc., 20(1), pp. 1–7. doi: 10.1016/j.intman.2013.10.003.

Ang, S. H., Benischke, M. H. and Doh, J. P. (2015) 'The interactions of institutions on foreign market entry mode', *Strategic Management Journal*, 36(10), pp. 1536–1553. doi: 10.1002/smj.2295.

Azar, G. and Drogendijk, R. (2018) 'How do Managers' Deviant Perceptions of "Cultural Distance" Relate to the Performance of International SMEs?', in Verbeke, A., Tulder, R. van, and Puck, J. (eds) *Distance in International Business: Concept, Cost and Value Vol: 12*. First. Bingley: Emerald Group Publishing LTD.

Beckerman, W. (1956) 'Distance and the Pattern of Intra-European Trade', *The Review of Economics and Statistics*, 38(1), pp. 31–40.

Belderbos, R., Du, H. S. and Slangen, A. (2020) 'When do firms choose global cities as foreign investment locations within countries? The roles of contextual distance, knowledge intensity, and target-country experience', *Journal of World Business*, 55(1). doi: 10.1016/j.jwb.2019.101022.

Berry, H., Guillén, M. F. and Zhou, N. (2010) 'An institutional approach to cross-national distance', *Journal of International Business Studies*, 41(9), pp. 1460–1480.

Beugelsdijk, S., Ambos, B. and Nell, P. C. (2018) 'Conceptualizing and measuring distance in international business research: Recurring questions and best practice guidelines', *Journal of International Business Studies*, 49(9), pp. 1113–1137. doi: 10.1057/s41267-018-0182-4.

Blanc-Brude, F. *et al.* (2014) 'The FDI location decision: Distance and the effects of spatial dependence', *International Business Review*, 23(4), pp. 797–810. doi: 10.1016/j.ibusrev.2013.12.002.

Buckley, P. J. *et al.* (2007) 'The determinants of Chinese foreign direct investment', *International Business*, 38(4), pp. 499–518.

Campos, J., Ericsson, N. R. and Hendry, D. F. (2005) 'General-to-specific Modeling : An Overview and Selected Bibliography', *International Finance Discussion Paper*, 2005(838), pp. 1–94. doi: 10.17016/ifdp.2005.838.

Child, J. (2000) 'Theorizing about organization cross-nationality: part 1 an introduction', in *Managing Across Cultures: Issues and Perspectives*. International Thomson Business Press. Available at: <https://books.google.pt/books?id=Ylu1AAAAIAAJ>.

Coase, R. (1937) 'The nature of the firm', *Economica*, 4, pp. 385–405. doi: 10.1016/S0039-6109(16)37642-3.

Corcoran, A. and Gillanders, R. (2014) 'Foreign direct investment and the ease of doing business', *Review of World Economics*, 151(1), pp. 103–126. doi: 10.1007/s10290-014-0194-5.

Dikova, D. (2009) 'Performance of foreign subsidiaries: Does psychic distance matter?', *International Business Review*, 18(1), pp. 38–49. doi: 10.1016/j.ibusrev.2008.11.001.

Dinner, I. M., Kushwaha, T. and Steenkamp, J. B. E. M. (2019) 'Psychic distance and performance of MNCs during marketing crises', *Journal of International Business Studies*. Palgrave Macmillan UK, 50(3), pp. 339–364. doi: 10.1057/s41267-018-0187-z.

Dow, D. (2018) 'Are We at a Turning Point for Distance Research in International Business', in Verbeke, A., Tulder, R. van, and Puck, J. (eds) *Distance in International Business: Concept,*

Cost and Value Vol: 12. First. Bingley: Emerald Group Publishing LTD.

Driscoll, A. M. and Paliwoda, S. J. (1997) 'Dimensionalizing international market entry mode choice', *Journal of Marketing Management*, 13(1–3), pp. 57–87. doi: 10.1080/0267257X.1997.9964459.

Duarte, M. P. and Carvalho, F. (2018) 'Institutional distance and multinational enterprises' FDI decisions in Portugal', *European Journal of Applied Business and Management*, 4(3), pp. 1–29.

Fayol, H. (1981) *Administração Industrial e Geral*. 9th edn. Edited by Atlas S.A. São Paulo.

Ghemawat, P (2001) 'Distance still matters - The hard reality of global expansion', *HARVARD BUSINESS REVIEW*, 79(8), p. 137+.

Ghemawat, P. and Hout, T. M. (2017) 'Administrative Distance and Institutional Variety', in *The Laws of Globalization and Business Applications*, pp. 280–320.

Guler, I. and Guillén, M. F. (2010) 'Institutions and the internationalization of US venture capital firms', *Journal of International Business Studies*, 41(2), pp. 185–205. doi: 10.1057/jibs.2009.35.

Haliti, B. *et al.* (2020) 'The impact of the Ease Doing Business Indicators on Foreign Direct Investment in the European transition economies', *Ekonomika*, 98(2), pp. 19–32. doi: 10.15388/ekon.2019.2.2.

Hasan, M. M., Ibrahim, Y. and Uddin, M. M. (2016) 'Institutional Distance Factors Influencing on Firm Performance: A Hypothetical Framework from Cross-Border Mergers and Acquisitions', *The Journal of Developing Areas*, 50(6), pp. 377–386. doi: 10.2139/ssrn.3042397.

Henisz, W. J. (2000) 'The Institutional Environment for Growth', *Economics and Politics*, 12(1), pp. 1–31.

Henisz, W. J. and Williamson, O. Y. (1999) 'Comparative Economic Organization-within and Between Countries', *Business and Politics*, 1(3), pp. 261–278. doi: 10.1515/bap.1999.1.3.261.

Hofstede, G. (1980) *Culture's consequences: International Differences in work-related values*. Beverly Hills, California: SAGE PUBLICATIONS INC.

Hofstede, G. (1991) *Cultures and Organizations: Software of the Mind*. New York, New York: McGraw-Hill.

Hofstede, G., Hofstede, G. J. and Minkov, M. (2010) *Cultures and organizations: software of the mind*. 3rd edn. New York, New York: McGraw-Hill.

Hofstede, G. and Minkov, M. (2010) 'Long- versus short-term orientation: New perspectives', *Asia Pacific Business Review*, 16(4), pp. 493–504. doi: 10.1080/13602381003637609.

Hossain, M. T. *et al.* (2018) 'Ease of Doing Business and Its Impact on Inward FDI', *Indonesian Journal of Management and Business Economics*, 1(1), p. 52. doi: 10.32455/ijmbe.v1i1.52.

House, R. J. *et al.* (2004) *Culture, Leadership and Organizations: The GLOBE Study of 62 Societies*. SAGE PUBLICATIONS INC.

Hutzschenreuter, T., Kleindienst, I. and Lange, S. (2016) 'The Concept of Distance in International Business Research: A Review and Research Agenda', *International Journal of Management Reviews*, 18(2), pp. 160–179. doi: 10.1111/ijmr.12065.

Jackson, G. and Deeg, R. (2008) 'Comparing capitalisms: Understanding institutional diversity and its implications for international business', *Journal of International Business Studies*, 39(4), pp. 540–561. doi: 10.1057/palgrave.jibs.8400375.

Johanson, J. and Vahlne, J.-E. (1977) 'The Internationalization Process of the Firm—A Model of Knowledge Development and Increasing Foreign Market Commitments', *Journal of International Business Studies*, 8(1), pp. 23–32. doi: 10.1057/palgrave.jibs.8490676.

Johanson, J. and Vahlne, J. E. (2009) 'The Uppsala internationalization process model revisited: From liability of foreignness to liability of outsidership', *Journal of International Business Studies*. Palgrave Macmillan, 40(9), pp. 1411–1431. doi: 10.1057/jibs.2009.24.

Johanson, J. and Wiedersheim-Paul, F. (1975) 'The Internationalization of the Firm — Four Swedish Cases', *Journal of Management Studies*, 12(3), pp. 305–323. doi: 10.1111/j.1467-

6486.1975.tb00514.x.

Kahneman, D. (2012) *Pensar, Depressa e Devagar*. Temas e De. Lisboa.

Kang, J., Lee, J. Y. and Ghauri, P. N. (2017) 'The Interplay of Mahalanobis Distance and Firm Capabilities on MNC Subsidiary Exits from Host Countries', *Management International Review*. Springer Berlin Heidelberg, 57(3), pp. 379–409. doi: 10.1007/s11575-016-0287-6.

Kaufmann, D., Kraay, A. and Mastruzzi, M. (2011) 'The worldwide governance indicators: Methodology and analytical issues', *Hague Journal on the Rule of Law*, 3(2), pp. 220–246. doi: 10.1017/S1876404511200046.

Kawai, N. and Chung, C. (2019) 'Expatriate utilization , subsidiary knowledge creation and performance : The moderating role of subsidiary strategic context', *Journal of World Business*. Elsevier, 54(1), pp. 24–36. doi: 10.1016/j.jwb.2018.09.003.

Knoll, S. *et al.* (2018) 'The role of distance on international commerce: the impact on Sino-Brazilian beef trade', *Journal of Agribusiness in Developing and Emerging Economies*, 8(4), pp. 693–711. doi: 10.1108/JADEE-04-2017-0041.

Kogut, B. and Singh, H. (1988) 'The Effect of National Culture on the Choice of Entry Mode', *Journal of International Business Studies*, 19(3), pp. 411–432.

Konara, P. and Shirodkar, V. (2018) 'The Direction of Regulatory Institutional Distance and MNE's Subsidiary Ownership Strategy: Re-examining Theory and Evidence in the Case of Emerging Markets', in Verbeke, A., Tulder, R. Van, and Puck, J. (eds) *Distance in International Business: Concept, Cost and Value Vol: 12*. First. Bingley: Emerald Group Publishing LTD.

Kostova, T. *et al.* (2019) 'The construct of institutional distance through the lens of different institutional perspectives: Review, analysis, and recommendations', *Journal of International Business Studies*. Palgrave Macmillan UK. doi: 10.1057/s41267-019-00294-w.

Kuo, C. L. and Fang, W. C. (2009) 'Psychic distance and FDI location choice: Empirical examination of Taiwanese firms in China', *Asia Pacific Management Review*, 14(1), pp. 85–106. doi: 10.6126/APMR.2009.14.1.06.

- Liu, K. *et al.* (2019) 'Foreign IPOs in the US: When entrepreneurial orientation meets institutional distance', *Journal of Business Research*, 101, pp. 144–151. doi: 10.1016/j.jbusres.2019.04.011.
- López-Duarte, C., Vidal-Suárez, M. M. and González-Díaz, B. (2016) 'International Business and National Culture : A Literature Review and Research Agenda', *International Journal of Management Reviews*, 18, pp. 397–416. doi: 10.1111/ijmr.12070.
- López-Duarte, C., Vidal-Suárez, M. M. and González-Díaz, B. (2019) *Cross-national distance and international business: an analysis of the most influential recent models*, *Scientometrics*. Springer International Publishing. doi: 10.1007/s11192-019-03203-y.
- Lubatkin, M. *et al.* (1998) 'Managing Mergers Across Borders : A Two-Nation Exploration of a Nationally Bound Administrative Heritage', *Organization Science*, 9(6), pp. 670–684.
- Luo, Y. (2002) 'Building trust in cross-cultural collaborations: Toward a contingency perspective', *Journal of Management*, 28(5), pp. 669–694. doi: 10.1016/S0149-2063(02)00161-7.
- De Maesschalck, R., Jouan-Rimbaud, D. and Massart, D. L. (2000) 'The Mahalanobis distance', *Chemometrics and Intelligent Laboratory Systems*, 50(1), pp. 1–18. doi: 10.1016/S0169-7439(99)00047-7.
- Mahalanobis, P. C., Bose, R. C. and Roy, S. Na. (1937) 'Normalisation of Statistical Variates and the Use of Rectangular Co-Ordinates in the Theory of Sampling Distributions', *Indian Statistical Insitutte*, 3(1), pp. 1–40.
- Malhotra, S., Sivakumar, K. and Zhu, P. (2009) 'Distance factors and target market selection: the moderating effect of market potential', *International Marketing Review*, 26(6), pp. 651–673. doi: 10.1108/02651330911001332.
- Maseland, R. (2018) 'Institutional Overlap as a Basis for International Business', in Verbeke, A., Tulder, R. van, and Puck, J. (eds) *Distance in International Business: Concept, Cost and Value Vol: 12*. First. Bingley: Emerald Groud Publishing LTD.

- Mingo, S., Morales, F. and Dau, L. A. (2018) 'The interplay of national distances and regional networks: Private equity investments in emerging markets', *Journal of International Business Studies*, 49(3), pp. 371–386. doi: 10.1057/s41267-017-0141-5.
- Minkov, M. and Hofstede, G. (2011) 'The evolution of Hofstede's doctrine', *Cross Cultural Management: An International Journal*, 18(1), pp. 10–20. doi: 10.1108/135276011111104269.
- Mueller, M. J., Hendriks, G. and Slangen, A. H. L. (2018) 'How the Direction of Institutional Distance Influences Foreign Entry Mode Choices: an Information Economics Perspectives', in Verbeke, A., Tulder, R. van, and Puck, J. (eds) *Distance in International Business: Concept, Cost and Value Vol: 12*. First. Bingley: Emerald Group Publishing LTD.
- Nau, H. H. (2005) 'Institutional, Evolutionary and Cultural Aspects in Max Weber's Social Economics', *Papers in Political Economy*, 47, pp. 127–142.
- Nelson, R. R. and Rosenberg, N. (1993) *National Innovation Systems: A Comparative Analysis*. Oxford University Press.
- North, D. C. (1991) 'Institutions', *Journal of Economic Perspectives*, 5(1), pp. 97–112.
- O'Grady, S. and Lane, H. W. (1996) 'The Psychic Distance Paradox', *Journal of International Business Studies*, 27(2), pp. 309–333. doi: 10.1017/CBO9781107415324.004.
- Ojala, A. and Tyrväinen, P. (2007) 'Market entry and priority of small and medium-sized enterprises in the software industry: An empirical analysis of cultural distance, geographic distance, and market size', *Journal of International Marketing*, 15(3), pp. 123–149. doi: 10.1509/jimk.15.3.123.
- Pajunen, K. (2008) 'Institutions and inflows of foreign direct investment: A fuzzy-set analysis', *Journal of International Business Studies*, 39(4), pp. 652–669. doi: 10.1057/palgrave.jibs.8400371.
- Parietti, L. V. (2017) 'Institutional Distance: Conceptualization and Measurements', in *Internationalization of Firms: The Role of Institutional Distance on Location and Entry mode*, pp. 1–37. doi: 10.1108/9781787141346.

Pinheiro-Alves, R. and Zambujal-Oliveira, J. (2012) 'The Ease of Doing Business Index as a tool for investment location decisions', *Economics Letters*. Elsevier B.V., 117(1), pp. 66–70. doi: 10.1016/j.econlet.2012.04.026.

La Porta, R. *et al.* (1998) 'Law and Finance', *Journal of Political Economy*, 106(6), pp. 1113–1155.

Qian, X. and Sandoval-Hernandez, J. (2016) 'Corruption distance and foreign direct investment', *Emerging Markets Finance and Trade*, 52(2), pp. 400–419. doi: 10.1080/1540496X.2015.1047301.

Quer, D., Claver, E. and Rienda, L. (2017) 'Cultural distance, political risk and location decisions of emerging-market multinationals: a comparison between Chinese and Indian firms', *Journal of the Asia Pacific Economy*, 22(4), pp. 587–603. doi: 10.1080/13547860.2017.1312065.

Reimann, F., Rauer, J. and Kaufmann, L. (2015) 'MNE Subsidiaries ' Strategic Commitment to CSR in Emerging Economies : The Role of Administrative Distance , Subsidiary Size , and Experience in the Host Country', *Journal of Business Ethics*. Springer Netherlands, 132, pp. 845–857. doi: 10.1007/s10551-014-2334-1.

Schwartz, S. H. (1994) *Beyond Individualism/Collectivism: New Cultural Dimensions of Values*. Thousand Oaks, California: SAGE PUBLICATIONS INC.

Scott, W. R. (2014) *Institutions and Organizations*. Thousand Oaks, California.

Shahadan, F., Sarmidi, T. and Jan Faizi, F. (2014) 'Relationships between Doing Business Indexes and FDI Net Inflows : Empirical Evidence from Six Asian Countries', *Persidangan Kebangsaan Ekonomi Malaysia*, 9, pp. 609–625.

Shenkar, O. (2001) 'Cultural Distance Revisited: Towards a More Rigorous Conceptualization and Measurement of Cultural Differences', *Journal of International Business Studies*, 32(3), pp. 519–535.

Shenkar, O. (2012) 'Beyond cultural distance : Switching to a friction lens in the study of

cultural differences', *Journal of International Business Studies*. Nature Publishing Group, 43, pp. 12–17. doi: 10.1057/jibs.2011.42.

Solberg, C. A. (2008) 'Product complexity and cultural distance effects on managing international distributor relationships: A contingency approach', *Journal of International Marketing*, 16(3), pp. 57–83. doi: 10.1509/jimk.16.3.57.

UNCTAD (2017) *Handbook of Statistics*. Geneva. Available at: https://unctad.org/en/PublicationsLibrary/tdstat42_en.pdf.

Vaccarini, K. *et al.* (2017) 'Chinese FDI and psychic distance perceptions on regulations in the German renewable energy sector', *ENERGY POLICY*, 101, pp. 723–732. doi: 10.1016/j.enpol.2016.09.008.

Verbeke, A., Puck, J. and Tulder, R. van (eds) (2018) 'When Distance is Good: An Upper-Echelons Perspective on the Role of Distance in Internationalization', in *Distance in International Business: Concept, Cost and Value Vol: 12*. First. Bingley: Emerald Group Publishing LTD.

Vlaj, D. *et al.* (2019) 'Expatriates managers' cultural intelligence as promoter of knowledge transfer in multinational companies', 94(February 2018), pp. 367–377. doi: 10.1016/j.jbusres.2018.01.033.

Whitley, R. (1992) *Business Systems in East Asia*. London: SAGE Publications INC.

Williamson, O. E. (1975) 'Markets and Hierarchies: Analysis and Antitrust Implications: A Study in the Economics of Internal Organization', *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

Williamson, O. E. (1985) *The Economic Institutions of Capitalism*. London: Collier Maxmillan Publishers

Wooldridge, J. M. (2013) *Introductory Econometrics, A Modern Approach*. 5th edn, *Applied Discrete-Choice Modelling*. 5th edn. Mason, Ohio: South-Western. doi: 10.4324/9781351140768-8.

World Bank Group (2020) *Doing Business 2020 - Comparing Business Regulation in 190 Economies*. Available at:

<https://openknowledge.worldbank.org/bitstream/handle/10986/32436/9781464814402.pdf>

.

Zaheer, S., Schomaker, M. S. and Nachum, L. (2012) 'Distance without direction : Restoring credibility to a much-loved construct', *Journal of International Business Studies*. Nature Publishing Group, 43(1), pp. 18–27. doi: 10.1057/jibs.2011.43.

8 – Appendix

Source of all the tables in the appendix: Author.

Table 10 – Baseline Study Model 1 – Summary Statistics

	N	Mean	Median	S.D.	Min	Max	VIF
FDI Deflated	268	2598	227.7	5004	-196.9	24964	
Admin_Dist	272	43.53	34.26	35.75	0.06129	142.1	3.512
Cultural_Dist	233	20.21	19.41	9.607	3.024	58.17	2.102
Demographic_Dist	272	5.900	2.658	7.013	0.1799	29.35	1.905
Economic_Dist	272	5.098	2.529	7.766	0.3992	49.14	1.756
Finance_Dist	199	3.755	2.409	3.418	0.07231	15.78	1.363
Geographic_Dist	272	4827	2930	4402	346.8	19801	2.434
Connect_Dist	264	2.465	2.064	2.573	0.03006	17.75	2.697
Know_Dist	256	4.439	2.297	6.461	0.001706	44.99	2.422
Political_Dist	272	156.0	124.3	61.02	57.21	235.1	2.777
Lang_PT	272	0.088	0	0.284	0	1	2.314
XR	272	58.25	1.740	248.3	0.4279	1753	1.242

Table 11 Baseline Study Model 1 – Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11	12
1 FDI_Deflated	1.0000	-.4061	-.0491	-.2807	0.0457	0.0237	-.3018	-.1183	-.1409	-.1930	-.1148	-.1003
2 Admin_Dist		1.0000	0.4025	-.0113	-.0665	-.1714	0.1168	0.6310	0.3119	0.2938	0.5097	0.0099
3 Cultural_Dist			1.0000	-.1267	-.2348	-.1747	0.1428	0.3284	0.5000	0.2749	0.1360	-.0438
4 Demographic_Di				1.0000	-.1946	0.2998	0.3013	-.0092	-.1226	0.2386	0.1468	0.3070
5 Economic_Dist					1.0000	0.1889	-.2088	0.0665	0.0638	-.0142	-.1066	-.0944
6 Finance_Dist						1.0000	0.0254	0.0072	-.1823	0.1018	-.1800	0.1235
7 Geographic_Dist							1.0000	0.0109	0.1757	0.5038	0.2395	0.1601
8 Connect_Dist								1.0000	0.0527	0.2511	0.0604	-.0440
9 Know_Dist									1.0000	0.3844	0.1403	-.0646
10 Political_Dist										1.0000	0.1597	-.0509
11 XR											1.0000	-.0384
12 Lang_PT												1.0000

Table 12 – Baseline Study Model 1a - GLS Regression

Model 1: Random-effects (GLS), using 179 observations
 Included 25 cross-sectional units
 Time-series length: minimum 1, maximum 8
 Dependent variable: FDI_Deﬂated
 Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	1917.32	1356.69	1.413	0.1576	
Admin_Dist	-113.734	35.3821	-3.214	0.0013	***
Cultural_Dist	100.220	31.7591	3.156	0.0016	***
Demographic_Dist	-17.2850	140.714	-0.1228	0.9022	
Economic_Dist	-0.815865	57.8894	-0.01409	0.9888	
Finance_Dist	-31.9809	90.5611	-0.3531	0.7240	
Geographic_Dist	-0.621810	0.361400	-1.721	0.0853	*
Connect_Dist	1021.33	401.044	2.547	0.0109	**
Know_Dist	-42.6248	50.6615	-0.8414	0.4001	
Political_Dist	31.4522	23.7650	1.323	0.1857	
XR	5.02965	2.35138	2.139	0.0324	**
Lang_PT	921.957	1698.34	0.5429	0.5872	

Mean dependent var	3821.543	S.D. dependent var	5743.927
Sum squared resid	3.74e+09	S.E. of regression	4716.926
Log-likelihood	-1762.459	Akaike criterion	3548.918
Schwarz criterion	3587.167	Hannan-Quinn	3564.428
rho	0.449535	Durbin-Watson	0.904471

Table 13 – Baseline Study Model 1b - GLS Regression

Model 2: Random-effects (GLS), using 157 observations
 Included 24 cross-sectional units
 Time-series length: minimum 2, maximum 7
 Dependent variable: FDI_Deﬂated
 Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	2477.20	1327.07	1.867	0.0619	*
Admin_Dist	-127.903	47.1268	-2.714	0.0066	***
Cultural_Dist_1	139.894	33.0873	4.228	<0.0001	***
Demographic_Dist_1	3.98479	131.282	0.03035	0.9758	
Economic_Dist_1	21.0295	66.4522	0.3165	0.7517	
Finance_Dist_1	-86.5927	77.5878	-1.116	0.2644	
Geographic_Dist	-0.698427	0.352995	-1.979	0.0479	**
Connect_Dist_1	1150.39	945.281	1.217	0.2236	
Know_Dist_1	-123.222	91.6333	-1.345	0.1787	
Political_Dist_1	32.2432	20.8683	1.545	0.1223	
XR_1	5.82622	2.77759	2.098	0.0359	**
Lang_PT	862.229	1776.83	0.4853	0.6275	

Mean dependent var	3947.880	S.D. dependent var	5943.744
Sum squared resid	3.38e+09	S.E. of regression	4814.376
Log-likelihood	-1548.331	Akaike criterion	3120.662
Schwarz criterion	3157.337	Hannan-Quinn	3135.557
rho	0.314538	Durbin-Watson	1.146316

Table 14 – Baseline Study Model 2 – Summary Statistics

	Mean	Median	S.D.	Min	Max	VIF
FDIStock_DeFlated	3036	449.1	6058	-337.4	33640	
Admin_Dist	43.53	34.26	35.73	0.06129	142.1	3.772
Cult_Mahal	1.726	1.912	1.069	0.1663	4.323	1.534
Demographic_Dist	6.699	2.865	8.158	0.1688	39.60	1.745
Economic_Dist	5.413	2.677	8.349	0.1207	60.39	1.351
Finance_Dist	3.993	2.720	3.498	0.07231	15.78	1.608
Geographic_Dist	4827	2930	4399	346.8	19801	2.035
Connect_Dist	2.255	1.741	2.267	0.03006	17.75	2.232
Know_Dist	6.489	2.882	10.78	0.001706	71.43	1.825
Political_Dist	160.9	130.0	60.40	57.21	238.4	2.625
Lang_PT	0.08824	0.0000	0.2840	0.0000	1.000	1.273
XR	58.26	1.674	243.7	0.4279	1753	2.638

Table 15 – Baseline Study Model 2 – Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11	12
1 FDIStock_Defla~	1.0000	-.4057	-.0983	-.2523	.1201	-.0995	-.3156	-.1165	-.1488	-.2342	-.0861	-.1103
2 Admin_Dist		1.0000	.2084	-.0186	-.0730	-.1605	.1168	.5820	.3168	.3142	.0099	.5160
3 Cult_Mahal			1.0000	-.3070	.1345	-.1934	.0205	.3972	-.0571	-.0381	-.2174	-.2796
4 Demographic_Di~				1.0000	-.1808	.3434	.2883	-.0154	-.0004	.2368	.3049	.1269
5 Economic_Dist					1.0000	.0606	-.2000	.0895	.0060	-.0286	-.0796	-.1109
6 Finance_Dist						1.0000	.0096	.0407	-.2008	.0626	.0520	-.1805
7 Geographic_Dist							1.0000	.0401	.2139	.4974	.1601	.2393
8 Connect_Dist								1.0000	-.0427	.2015	.0495	.0526
9 Know_Dist									1.0000	.3679	-.0406	.2189
10 Political_Dist										1.0000	-.0284	.1856
11 Lang_PT											1.0000	-.0375
12 XR												1.0000

Table 16 – Baseline Study Model 2a - GLS Regression

Model 1: Random-effects (GLS), using 296 observations
 Included 25 cross-sectional units
 Time-series length: minimum 2, maximum 13
 Dependent variable: FDIStock_Deflated
 Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	6367.22	3507.91	1.815	0.0695	*
Admin_Dist	-103.654	36.2125	-2.862	0.0042	***
Cult_Mahal	-62.3306	1117.87	-0.05576	0.9555	
Demographic_Dist	17.8645	135.202	0.1321	0.8949	
Economic_Dist	205.387	96.7109	2.124	0.0337	**
Finance_Dist	-512.016	235.066	-2.178	0.0294	**
Geographic_Dist	-0.374191	0.263809	-1.418	0.1561	
Connect_Dist	171.346	626.610	0.2734	0.7845	
Know_Dist	-21.1482	20.0228	-1.056	0.2909	
Political_Dist	20.4810	11.0679	1.850	0.0642	*
Lang_PT	594.403	1922.16	0.3092	0.7571	
XR	3.61133	1.72082	2.099	0.0359	**

Mean dependent var	4147.065	S.D. dependent var	6853.219
Sum squared resid	1.13e+10	S.E. of regression	6290.860
Log-likelihood	-3003.469	Akaike criterion	6030.939
Schwarz criterion	6075.223	Hannan-Quinn	6048.670
rho	0.742428	Durbin-Watson	0.421483

Table 17 – Baseline Study Model 2b - GLS Regression

Model 2: Random-effects (GLS), using 274 observations
 Included 25 cross-sectional units
 Time-series length: minimum 2, maximum 12
 Dependent variable: FDIStock_Deflated
 Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	6687.11	3692.39	1.811	0.0701	*
Admin_Dist	-113.679	38.0094	-2.991	0.0028	***
Cult_Mahal_1	-44.8008	1118.36	-0.04006	0.9680	
Demographic_Dist_1	58.9119	170.769	0.3450	0.7301	
Economic_Dist_1	199.777	104.135	1.918	0.0551	*
Finance_Dist_1	-560.957	273.622	-2.050	0.0404	**
Geographic_Dist	-0.406065	0.273474	-1.485	0.1376	
Connect_Dist_1	344.537	394.703	0.8729	0.3827	
Know_Dist_1	-25.2636	22.4570	-1.125	0.2606	
Political_Dist_1	20.5125	9.77714	2.098	0.0359	**
Lang_PT	469.252	2101.05	0.2233	0.8233	
XR_1	4.01723	1.73177	2.320	0.0204	**

Mean dependent var	4245.411	S.D. dependent var	7026.547
Sum squared resid	1.07e+10	S.E. of regression	6371.364
Log-likelihood	-2783.298	Akaike criterion	5590.595
Schwarz criterion	5633.953	Hannan-Quinn	5607.998
rho	0.715647	Durbin-Watson	0.482349

Table 18 – Model 1, with Vertical Administrative Distance addition – Summary Statistics

	Mean	Median	S.D.	Min
FDI_Deflated	3120	233.9	5721	-196.9
Hor_Admin_Dist	43.97	32.13	37.54	0.06129
Cultural_Dist	20.72	21.08	10.11	3.321
Demographic_Dist	5.893	2.700	7.313	0.1799
Economic_Dist	5.286	2.522	8.484	0.4096
Finance_Dist	4.054	2.578	3.538	0.07231
Geographic_Dist	4403	2786	4477	346.8
Connect_Dist	2.689	2.289	2.643	0.2465
Know_Dist	3.411	2.355	3.807	0.001706
Political_Dist	151.7	123.2	60.49	58.38
XR	63.21	1.667	268.6	0.5714
Vert_Admin_Dist	4.481	4.553	1.073	2.064

Table 19 – Reduced Model with Vertical Administrative Distance addition

Model 9: Random-effects (GLS), using 137 observations

Included 23 cross-sectional units

Time-series length: minimum 5, maximum 6

Dependent variable: IDE_Deflated

Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	0.412443	0.128009	3.222	0.0013	***
Vert_Admi_Dist	-0.205632	0.0892680	-2.304	0.0212	**
Hor_Admin_Dist	-0.489475	0.113358	-4.318	<0.0001	***
Cultural_Dist	0.165318	0.0715598	2.310	0.0209	**
Geographic_Dist	-0.262918	0.135935	-1.934	0.0531	*
Connect_Dist	0.107777	0.0476775	2.261	0.0238	**
Know_Dist	-0.167402	0.0609704	-2.746	0.0060	***
XR	0.343542	0.102063	3.366	0.0008	***

Mean dependent var	0.163195	S.D. dependent var	0.246193
Sum squared resid	4.741274	S.E. of regression	0.190975
Log-likelihood	36.01716	Akaike criterion	-56.03432
Schwarz criterion	-32.67447	Hannan-Quinn	-46.54144
rho	0.079922	Durbin-Watson	1.282167

'Between' variance = 0.0358427

'Within' variance = 0.00460706

mean theta = 0.854592

Table 20 – Model 5, with Starting a Business variables, not normalized

Model 5: Random-effects (GLS), using 102 observations					
Included 18 cross-sectional units					
Time-series length: minimum 1, maximum 6					
Dependent variable: IDE_Deflated					
Robust (HAC) standard errors					
	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	4517.75	3201.32	1.411	0.1582	
Admin_Dist	-113.687	52.3227	-2.173	0.0298	**
Cultural_Dist	-28.3657	61.2911	-0.4628	0.6435	
Demographic_Dist	-273.584	228.335	-1.198	0.2309	
Economic_Dist	-78.8826	62.6656	-1.259	0.2081	
Finance_Dist	111.643	122.543	0.9111	0.3623	
Geographic_Dist	-0.726286	0.480894	-1.510	0.1310	
Connect_Dist	500.647	456.502	1.097	0.2728	
Know_Dist	-568.740	222.066	-2.561	0.0104	**
Political_Dist	64.1361	41.2690	1.554	0.1202	
XR	8.53278	4.48861	1.901	0.0573	*
ProceduresMennumber	-183.745	186.745	-0.9839	0.3251	
TimeMendays	-35.3301	17.3341	-2.038	0.0415	**
CostMenofincomepercentage	189.167	112.403	1.683	0.0924	*
PaidinMinimumcapital	-17.8765	8.79495	-2.033	0.0421	**
Mean dependent var	5191.547	S.D. dependent var		6718.743	
Sum squared resid	2.95e+09	S.E. of regression		5790.103	
Log-likelihood	-1020.921	Akaike criterion		2071.841	
Schwarz criterion	2111.216	Hannan-Quinn		2087.785	
rho	0.000533	Durbin-Watson		1.502582	

Table 21 – Model 6, with Dealing with Construction Permits variables, not normalized

Included 18 cross-sectional units
Time-series length: minimum 1, maximum 6
Dependent variable: IDE_Deflated
Robust (HAC) standard errors

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
const	4246.43	3692.76	1.150	0.2502	
Admin_Dist	-133.169	49.9153	-2.668	0.0076	***
Cultural_Dist	63.6138	44.5422	1.428	0.1532	
Demographic_Dist	95.7030	221.240	0.4326	0.6653	
Economic_Dist	2.91630	36.2239	0.08051	0.9358	
Finance_Dist	34.9733	75.5697	0.4628	0.6435	
Geographic_Dist	-0.663676	0.487901	-1.360	0.1737	
Connect_Dist	583.794	282.030	2.070	0.0385	**
Know_Dist	-309.633	154.000	-2.011	0.0444	**
Political_Dist	49.0528	40.0364	1.225	0.2205	
XR	8.35465	4.34067	1.925	0.0543	*
Proceduresnumber	-143.482	171.595	-0.8362	0.4031	
Timedays	-20.2326	16.9731	-1.192	0.2332	
CostofWarehousevalue	1032.50	712.511	1.449	0.1473	
Mean dependent var	5191.547	S.D. dependent var		6718.743	
Sum squared resid	3.39e+09	S.E. of regression		6173.152	
Log-likelihood	-1028.031	Akaike criterion		2084.062	
Schwarz criterion	2120.812	Hannan-Quinn		2098.943	
rho	-0.049403	Durbin-Watson		1.590323	