Copenhagen Business School

Master Thesis

"Cashless Society: Cross-Country Comparison of Key Drivers"

Authors: Francesc Aubia Ferré and Bruno Rita Real Caldas Simões

Supervisor: Prof. Jonas Hedman

Student no.: 116715 and 116375

Number of Pages and Characters: 62 pages, 118.767 characters

Date of Submission: 15.05.2019

Abstract

In the past few decades, the Cashless Society has attracted the attention of some studies and a great deal of expectations has been raised in this regard. Regardless of the widespread literature addressing the topic, no relevant publications have tried to analyse the factors that contribute to making a country move towards a Cashless Society from an empirical and holistic perspective. The purpose of this thesis is to identify the main drivers that lead to a Cashless Society, as well as the interaction between these drivers and the resulting impact on the level of cashlessness. In order to solve this knowledge gap in the existing academia, a statistical model has been developed with the aim of validating the initial hypothesised drivers -based on previous research- and measuring their influence across countries.

Our findings suggest that the level of cashlessness in a country is influenced by 6 drivers: (1) Degree of Digitalization; (2) Digital Trust & Privacy Concerns; (3) Legal Framework; (4) Maturity of the Banking Industry; (5) Transparency & Corruption, and (6) Economic Development & Financial Inclusion. The outcomes of our analysis reveal that the Degree of Digitalization; Maturity of the Banking Industry and Economic Development & Financial Inclusion have a positive impact on the level of cashlessness. Conversely, Digital Trust & Privacy has a negative effect. As for the Legal Framework, it is a double-edged sword as previously suggested in the existing literature.

This study contributes to the understanding of the payments digitalization phenomenon and can inspire further research intended, for instance, to analyse the consequences of a Cashless Society.

Keywords: *Cashless Society; Cashlessness; Payments Digitalization; Innovation; Trust; Privacy; Transparency; Key Drivers; Regression Analysis; Country Clustering*

Acknowledgements

We would like to acknowledge our supervisor, Prof. Jonas Hedman, for the patient guidance and expert advice that have been invaluable throughout all stages of the thesis. His expertise in Cashless Societies and Digitalization was of great value to redirect our initial proposal and ensure that the resulting findings could make a contribution to the existing research.

We must also express our gratitude to Cristina Leal Rodríguez, Phd Student at Novo Nordisk Foundation Center for Protein Research, for her dedication and commitment in developing and applying the statistical model presented in this paper and her effort to get quickly familiar with concepts not related to her field of research. Her support was crucial for the success of this study.

We would also wish to thank Hans Henrik Hoffmeyer, former Vice President of Smart Payments (Nets Group), and Alexandre Rodrigues for their valuable suggestions, who have contributed greatly to the improvement of the thesis.

Finally, we want to dedicate this work to David Silva Melo, good friend and flatmate of one of the authors, who passed away by suicide during the course of this paper.

Table of Contents

1. Introduction

This introductory section will provide readers with a concise overview of the topic analysed and, in particular, with a precise description of the problem statement examined throughout this research. Furthermore, the explanation of the theoretical and quantitative delimitations constraining the research will be carefully exposed. Ultimately, the consistency of the findings together with their supportive contribution to the existing academia will be presented.

1.1. Research question

In social sciences, academia mainly addresses two types of problems, one whose aim is to increase knowledge and the other whose aim is to improve quality of life (Selamat, 2008). This thesis sheds light on an acknowledged concept that has attracted the attention of many academics over the past years (Fabris, 2019): Cashless Society. This puzzling concept of a society which no longer uses any physical currency (Akinola, 2012) has persuaded several academics to propose empirical evidence and theoretical frameworks in order to formulate a plausible explanation. There are several studies published that have tried to explain part of the phenomenon from multiple angles and perspectives. From Kenneth Rogoff (2016) who through his book provides intriguing insights on how and why cash should be phased out and Bátiz-Lazo and Efthymiou (2016) who provide different views on a Cashless Society from around the world, to studies with a more digital focus that assess the impact of innovation on cashlessness (Akinola, 2012; Hedman & Henningsson, 2015) or, for example, on trust (Achord, et al., 2017; OECD, 2002).

Indeed, despite the extensive literature addressing the topic as well as the on-going academic debates, no relevant publications have intended to analyse those factors that make a country move towards a Cashless Society from an empirical perspective. Therefore, it can be concluded that there is a potential knowledge gap that should be bridged and whose results can contribute to further research, for instance, by focusing on one of the factors identified or by analysing the positive or negative impact of these drivers on the resulting Cashless Society.

In light of this, the core research of this thesis is aimed at providing an answer to the following questions: "**What are the main drivers towards a Cashless Society?"** and "**To what extent do they affect the cashlessness of each country?"** More specifically, and based on the literature review, six hypotheses have been formulated as potential drivers:

1) Degree of Digitalization
2) Digital Trust & Privacy Concerns
3) Legal framework
4) Maturity of the banking industry
5) Transparency & Corruption
6) Economic Development & Financial inclusion

As for the resulting outcome -**the level of cashlessness**-, it has been analysed from a **static** and **dynamic perspective** by not only taking the country's level in a certain year, but also its corresponding evolution.

1.2. Delimitations

Cashless Society is a comprehensive concept that involves complex and interacting dynamics with different stakeholders, including national and international actors and regulations. Delimitations to the scope of this thesis have been thus reported; they turn to be critical and indispensable for a correct interpretation regarding the internal and external validity of the results achieved in this thesis.

Delimitations to the magnitude and degree of this thesis do exist, mainly due to the complexity of the concept and the number of potential existing drivers. There is some risk of omitted-variable bias and, additionally, some drivers might not be included in our statistical model. This study includes data from 2008 to 2015 gathered from credible and reliable sources. However, determining whether this primary data was properly collected or not is out of the scope of this paper. At the same time, it is plausible to test our model in the years to come by simply updating the variables provided or, alternatively, by considering additional drivers based on future academic findings or using different variables for the drivers already listed.

1.3. Contribution to literature

Many of academic literature dedicated to this topic focuses on the advantages and disadvantages of a future Cashless Society. Other scholars decided to focus solely on certain aspects of Cashless Societies such as the cost of cash or the diversity of digital transactions and analyse them in depth.

Our thesis has the potential to contribute to the understanding of the main drivers' role towards a Cashless Society by developing a statistical model that explains how the different drivers interact with each other and impact the overall level of cashlessness of a country. In this matter, the development of the statistical model, analysis and statistical significance criteria will be resolutely discussed in the following sections of this thesis.

Finally, it is important to emphasize that the resulting dataset for this study is unique and is founded on the most recent and relevant studies. Additionally, the presented results can be validated by simply updating the values of the variables thanks to the periodicity of publications chosen.

1.4. Personal considerations

We would like to highlight that our thesis does not intend, in any matter, to give an opinion whether the move towards a potential Cashless Society will change society and individuals' life for the better or worse.

Lastly, we honestly believe that the findings of our study may have legitimate implications for further research and for respective stakeholders.

2. Literature Review

In this section, a detailed description of an imminent Cashless Society and the preeminent drivers behind the implementation of such a critical step in a country's economy will be provided. Moreover, an explanation of the pivotal distinct drivers will be reported, along with an explanation of the crucial stakeholders participating in the process and the necessary pre and post Cashless Society stages to undertake. Further, the relevant theories and empirical evidence related to the Cashless Society concept provided by academia will be discussed and scrutinized in detail.

2.1. Cashless society

2.1.1. Definition of topic

In the recent years, cash has got greater attention, as central banks and other public institutions worry more about how far they can lower interest rates below zero, as security agencies attempt to control terrorist threats, as governments' national treasuries become progressively reckless for tax revenues, and as justice departments endeavour to reduce international and national crime syndicates (Rogoff, 2016). In the last decades, there has been much speculation about a cash-free society (Akinola, 2012). From the moment that the original general-purpose charge card appears in the middle of the last century, scholars have predicting the "cashless society" (Garcia-Swartz, Hahn, & Layne-Farrar, 2006). Successive generations of economists and researchers debated about the future of cash (Fabris, 2019) and theorized their use while anticipating their societies in a future where there was no purpose for material representations of money (Bátiz-Lazo & Efthymiou, 2016).

Cashless societies have existed since the moment mankind came into existence, based on barter and other methods of exchange. However, the true Cashless Society should be understood in the sense of a move towards, as well as the consequences of, a society in which cash is replaced by its digital corresponding. In other words, "legal tender (money) exists, is recorded, and is exchanged only in electronic digital form" (Fabris, 2019). According to Akinola (2012), "a cashless society is a community in which all payments are electronic"; a community in which everything is paid through digital electronic money, for instance with online payments, credit or debit cards or mobile payments. A society in which nobody uses cash (ibid).

Henceforth 1888, the novelist Edward Bellamy anticipated a cashless society by the turn of the new millennium, but it still hasn't happened utterly (Warwick, 2004). Undoubtedly, the latest technological disruptions have led to discussions of reaching a cashless society -or a society with a complete level of cashlessness- in the near future. However, it is fairly obvious that this is a change of form rather than substance. In sum, a society becomes cashless when it is possible to envisage a payment technology which makes no use of paper money or metal coins or, in other words, in which "cash is not a generally accepted means of payment" (Hedman, 2018). Despite the significant progress of certain countries towards turning cashless, a "moneyless" society is yet to come (Smithin, 2000).

2.1.2. Road to a Cashless Society and framing the change

Becoming digital is the future of money (OECD, 2002). Over time, there have been several forms of payment systems, most notably barter, gold, and paper currency (Garcia-Swartz, Hahn, & Layne-Farrar, 2006). Upon closer examination at the history of money, it can be observed that it went through evolutionary changes starting from the barter system, precious metal used as the means of payment, money made from precious metal and gold baked money to money whose value is completely separate from the material it was made (Fabris, 2019).

For a very long time, coins and notes were the only options to make purchases and to exchange money between organizations and individuals. The second half of the twentieth century observed new ways of thinking systematically about methods for retail payments (Bátiz-Lazo & Efthymiou, 2016). The introduction of plastic cards, pre-paid payment cards, Electronic Fund Transfer and Internet banking all aimed at making payment more convenient (Akinola, 2012).

An important step forward to a cashless society came with the arrival of electronic funds transfer (EFT) technology, which conducted in the era of credit-card transactions around 1960 (Warwick, 2004). If last century was the epoch of cash, checks and plastic credit cards, the twenty-first century is one of swiftly increasing forms of electronic value transfer systems, each operating on distinct platforms using various protocols and network infrastructures (Maurer, 2016). With the development in information technology, a number of new payment solutions have emerged, such as, among others, SMS payments, PayPal, M-PESA, Bitcoin, Google Wallet, and WeChat (Akinola, 2012).

Although the Cashless Society has not yet fully become a reality, payment options by merchants and consumers have been moving in that direction over the past five decades (Garcia-Swartz & Layne-Farrar, 2006). For several years, money has been in the direction of increased abstraction, either absolute symbolic representation detached from an explicit physical materialization (OECD, 2002). Historically, countries have witnessed a flourishing trend in cashless transactions as well as products and services sold solely in this way (Fabris, 2019). Digital alternatives to cash have been in existence for quite some time and have advanced with the financial sophistication of their users and the payment technologies (Achord, Chan, Nardani, & Rochemont, 2017). By and large, these innovations are modifying people 's perception of money and experience of paying (ibid). The innovation process in payments is encouraged by different stakeholders (such as financial institutions, FinTech's, merchants, mobile operators, etc.) and by international governmental agencies (Akinola, 2012). In addition, new payment solutions attract a lot of attention in the press and media, but they also trigger a debate regarding the cashless society (ibid).

Ultimately, it is a complex transformation where politics, laws, business interests, values, technologies, power games and habits play an essential part (Arvidsson, 2019). This complexity needs to be recognized when intending to comprehend the transformation process even if the complexity at the same time makes it difficult to identify what kind of components are the most critical ones and in which direction the process will unfold (ibid).

2.2. Main drivers towards a Cashless Society

As discussed earlier, the fact that people are making more digital transactions than ever is undeniable (see Figure 1 below). In some countries, such as Sweden or Canada an imminent Cashless Society is around the corner (Smith, 2017). For others, there is still a long journey and a lot of ground to cover.

Regardless of the stage of cashlessness, after extensive research through the available and accessible literature review, it was possible to identify the main common drivers that lead towards a Cashless Society. Nevertheless, it is important to consider that for each country each driver will have a different level of interaction or intensity and therefore different outcomes.

This section is, therefore, going to provide evidences that support the choice of certain drivers as the ones with the highest impact on the road to Cashless Society.

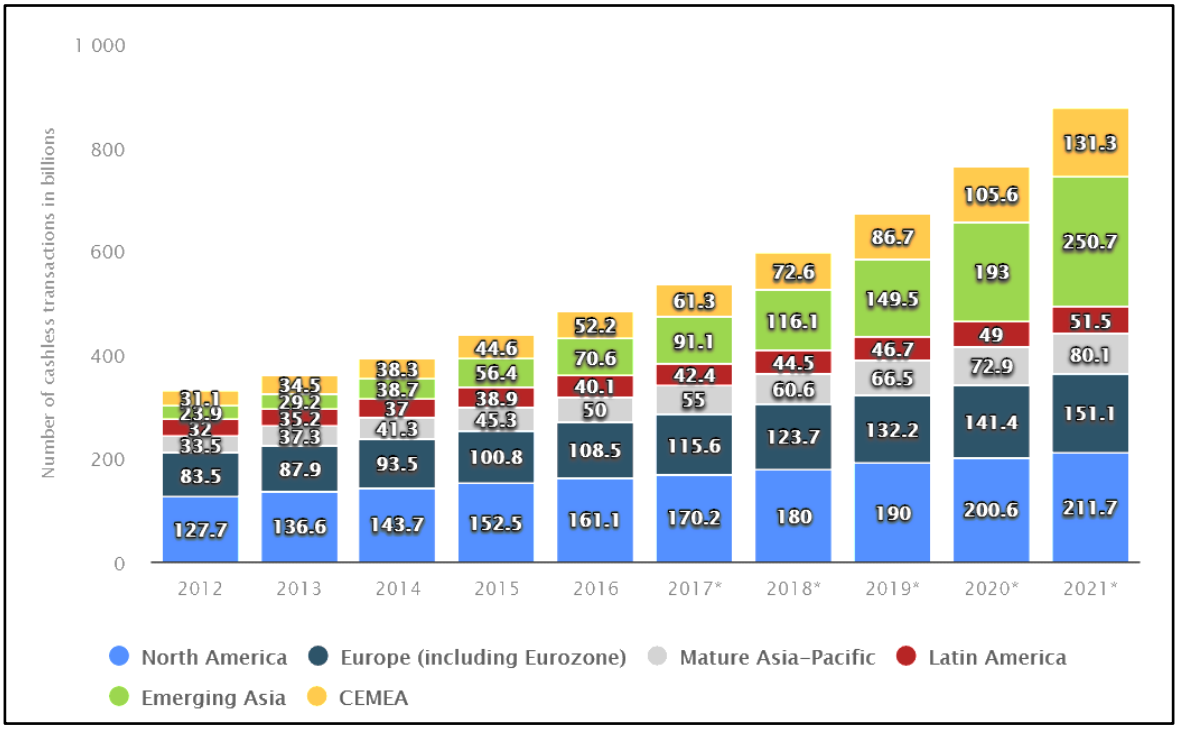


Figure 1: Number of non-cash transactions worldwide from 2012 to 2021, by region (in billions) – Statista 2019

2.2.1. Degree of Digitalization

One of the main drivers in the fast-changing world of payments is innovation triggered by new technological solutions. According to Fabris (2019), the developments of technology in the digital society lead to an increase in digital payments. Before, cards were exclusively used for high-value transactions, while cash was widely used for low-value transactions (Achord, Chan, Nardani, & Rochemont, 2017). Arvidsson (2019) argues that new payment services and technological solutions that have a purpose akin to cash are substituting cash in payment situations in which cash used to be the main means of payment.

Currently, the development of new digitized techniques for registering transactions and debiting accounts has been the main pathway for innovation (Baubeau, 2016). Frictionless, efficient payments such as NFC (Near-field Communication), QR codes and mobile payments made the clearance, the transfer of money from one party to another and the settlement of transactions more secure, reliable and efficient (Maurer, 2016). Cashless technologies provide worldwide payment networks and business solutions; simultaneously, payment technologies are unceasingly changing (Bátiz-Lazo & Efthymiou, 2016). Here again, the success of dematerialized money rests on strong material factors

such as traditional infrastructures (ibid). In this respect, as virtual as money may become, it will ultimately rely and depend upon tangible equipment and costly know-how (Baubeau, 2016).

Notwithstanding, what may be likely in the near future and cannot be excluded is the advent of digital currencies under the control of central government, at the very least in certain countries such as Sweden and as a replacement to cash (Fabris, 2019). Digital currency is a form of electronic money combined with new technology involving cryptography, peer-to-peer networking, databases and a system of consensus (Achord, Chan, Nardani, & Rochemont, 2017). The most remarkable example of digital currency is bitcoin. This also supported by the fact that a number of central banks are very actively investigating the costs and benefits of introducing these new forms of money (Fabris, 2019).

Internet has transformed transactions, especially in countries with poor financial infrastructures (Baubeau, 2016). For Africa's communication companies, this comes as an outstanding opportunity (as occurred with M-PESA), but to incumbent and established banks and supervision authorities, it is associated with high costs and risks (Akinola, 2012). And newcomers can take advantage of these new technological innovations to invent new forms of money, such as cryptocurrencies (ibid). The more alternative payment methods there are, the more competition and the lower the costs for merchants and consumers who want to make and receive payments at the expense of the banks (Achord, Chan, Nardani, & Rochemont, 2017).

Convenience plays a major part in the lives of most individuals in today's society. "The level of ease that the payer experiences when using the payment instrument depends on how many barriers or how much effort the user must make when completing a payment" (Holst, Hedman, Kjeldsen, & Tan, 2015). The constantly growing and expanding market of mobile devices in today's fast advancing world has initiated numerous innovative functions and services that are available on mobile devices (Hack, 2016). The use of mobile technology greatly reduces the cost of sending money over large distances, provides certainty of process and decreases the risk of theft (Rouse & Verhoef, 2016).

Mobile payment systems are the applications of mobile devices that are expected to grow robustly in the near future (Hack, 2016). This is not valid only for developed countries, but also or even more in developing countries. For example, only one in every four persons has a bank account in Africa, but four in five have access to a mobile phone (Rouse & Verhoef, 2016). In 2015, around 200 million people have access in Africa to the Internet due to their mobile phones (Afi, 2013). Mobile payment

systems provide banking services to the majority of unbanked inhabitants of many developing countries(ibid). This is largely as a result to the fact that mobile payment systems do not require exceedingly costly tangible assets and know-how (Fabris, 2019). Undoubtedly, mobile payments are one of many technological innovations that are disrupting the payment market (Hedman & Henningsson, 2015).

2.2.2. Trust in Digital Technologies & Privacy Concerns

According to Achord (2017) and Smithin (2000), another relevant driver of a cashless society is trust or more specifically, trust in digital technologies. Legitimate and stable political authority usually goes hand in hand with the level of digital money present in a determinate country (OECD, 2002). Although the majority of the end customer embrace the newer payment options (Achord, Chan, Nardani, & Rochemont, 2017), there must be trust between the payee and payer, faith in the regulatory setting for consumer protection, confidence in the safety and protection of the payment as well as conviction that the procedure is advantageous (Kapron & Meertens, 2017). “Trust is the foundation to any payment and is primarily concerned with ensuring that the payment credentials are handed over to the actual receiver of a payment” (Holst, Hedman, Kjeldsen, & Tan, 2015) .

Money as a store of abstract value consists in the social system of monetary production which entails the creation of monetary legitimacy which is a form of impersonal trust (Shapiro, 1987). Cashless Society requires that each and every of society’s stakeholders -individuals, governments, financial institutions and organizations- agree on and trust in digital currency (Akinola, 2012). Credit organizations and banks, traditional providers of banking services, have a clear competitive advantage over new entrants to this field (Hack, 2016). From a competitive point of view, it will be very interesting to see how banks can build on that advantage or whether technology companies such as Apple will provide enough advantages in terms of usefulness and ease of use to outweigh the lead of banks in terms of trust (ibid).

The anonymity, untraceability and decentralization of cash empower freedom to citizens against an omnipotent government or central bank (Achord, Chan, Nardani, & Rochemont, 2017). Most public leaders and scholar are cautious when considering the idea of an electronic currency system replacing cash in the near future, especially one regulated by the government, for fear that they are seen as amenable to compromise privacy (Warwick, 2004).

One key inhibitor for the level of trust in digital technologies and therefore for cashlessness amongst society could be the risk of a "complete loss of privacy" for individuals and the long-debated dilemma of who will have access to the supervision of their data (Fabris, 2019). When all personal and confidential information are vulnerable to the state, people can assume that the government has considerable authority over people by possessing access to such types of private data (Akinola, 2012).

In the context of rising uncertainty and disruption, "self-fulfilling long-term trust" is based on political and social legitimacy through which conceivably unreliable "strangers" are capable to take part in "impersonal complex multilateral economic relationships" (Smithin, 2000). A number of people appreciate to do certain purchases using cash so that their privacy is preserved, with no electronic record or audit report over such transactions (Brown, 1997). Despite the fact that literature often finds arguments that privacy is only required by individuals who have something to hide, this does not have to be the case (Fabris, 2019) . For instance, this information makes it possible to customer profiling, use of personal information for commercial reasons, the development of databases about their consumer habits, as well as profound knowledge into their belongings which can increases the risk of robbery, and so forth (ibid). One idea that can be implemented in order to avoid privacy concerns amongst individuals is to allow relatively small expenditures, up to a few hundred dollars or similar, to be confidential, even from the government (Rogoff, 2016). Essentially, the degree of trust in digital technologies is a kind of trade-off between privacy and convenience.

2.2.3. Legal Framework

Most scholars (Akinola, 2012) agree that a cashless society or a plan for phasing out most paper currency can only be implemented by the government or a central bank since they are the responsible for printing money and also control the supply of cash in society (Rogoff, 2016). The Organization for Economic Cooperation and Development (2002), argues that policymakers are responsible and have good reasons to pick up the pace at which the digital money diffuses throughout the economy. Given that the cash payments market is highly regulated (Arvidsson, 2019), a Cashless Society will only be possible if a government commit itself to the project, taking into account that just the government can actually control an electronic replacement for cash. Moreover, it may place an end to the circulation and production of cash (Warwick, 2004).

Some governments encourage a shift toward a Cashless Society because they see it as a way to avoid tax evasion and address money laundering in addition to boosting competition in financial services (Fabris, 2019). Most of this encouragement is made through the form of rules and compliance legislation (ibid). Cyber security protection, regarding companies and also individuals, is one of the greatest concerns contemplated in the discussion when elaborating new regulations (Achord, Chan, Nardani, & Rochemont, 2017). Legal frameworks and regulatory incentives intent to discourage other means of making substantial payments that can be altogether hidden from the government (Rogoff, 2016).

In most occasions, cashless technological innovations cannot be implemented due to a lack of appropriate legislation, while on other few occasions, a lack of a clearly defined regulatory framework had the opposite effect of incentivizing trade and commerce to take the initiative and develop its own cash- less payment instruments (Bátiz-Lazo & Efthymiou, 2016). Nevertheless, since phasing out paper currency is an exceptionally extensive concept that involves complex and interacting dynamics with different stakeholders (Arvidsson, 2019), national and international regulation should be put in (Rogoff, 2016). This process should be done gradually in order to avoid excessive disruption and to give individuals and institutions time to adapt (ibid). Policymakers should introduce regulatory incentives, initiatives and rules, as well as nurturing the institutions that run complex settlement systems (OECD, 2002). These kinds of regulatory incentives could be able to transform cashlessness into a more pragmatic and efficient economic reality (ibid).

According to Blind (2015), policy uncertainty, but also compliance regulation do appear to cause both negative and positive effects. These effects can be differentiated between short- and long-term impacts. The negative effects of compliance costs as well as regulatory ineffectiveness provoked by uncertainty or weak institutions are most relevant in the short run. This short-term impact is compared with the more dynamic effect of regulations generating additional incentives for innovative activities. Whereas, in the short term, the required regulatory compliance creates a burden for most companies, which might be negative for innovation, new legal frameworks that regulate the implementation of innovative technologies or solutions can incentive their development by providing compliance guidance and legal certainty. The net impact of regulation on innovation depends on the extent of the policy uncertainty and compliance cost on the one hand and the incentive effect on the other hand.

2.2.4. Maturity of the Banking Industry

It is important to state that a regulatory framework governing the usage of cash can only be implemented if it is complemented by a "decentralized, operative and market-driven structure" (Arvidsson, 2019) composed by actors such as ATMs, banks or merchants, among others. Banks, as part of this complex structure, bear a deeper business interest in a transition to a cash-free society since maintaining supply cash handling services became costlier and represents no opportunity for cross-selling related products linked to cash (ibid).

The prevailing acceptance of debit and credit payment cards in the middle of the twentieth century, the digitalization of bank accounts in the mid-1960s and the establishment of Internet and electronic banking around the 90s (Hedman, 2018); contribute to the maturity of the banking system. Digital usage has become firmly associated with customer loyalty with banks (Robin, 2015), regardless of relinquish the anonymity of paper currency with non-anonymous electronic money (Kenneth, 2015).

Inequality between countries and amongst the rich and poor within them remain partly due to the regulation of retail financial markets, custom and culture (Batiz-Lazo, Efthymiou, & Michael, 2016). Rouse and Verhoef (2016) state that in most developed countries, banks and other financial institutions are increasingly making the move from human to digital banking and therefore shaping the transition to a cashless society. By making a priority integrated money management, the flexibility of use and ease of access to banking services, payment mechanisms and monitoring platforms (ibid).

According to Afi (2013), in the absence of well-established bank networks in the more developing countries, most people bear preference for cash and tend to display a distrust or scepticism towards mobile money transfer mechanisms. A payment service has to connect to the same technological platform as many supplier and users to be beneficial for the society as a whole (Arvidsson, 2019), therefore the importance of the development of the banking industry.

2.2.5. Transparency & Corruption

The main issue of having a cash-free society is whether the benefits would outweigh disadvantages. (Akinola, 2012). According to the renowned economist Kenneth S. Rogoff (2016), the ultimate goal to phasing out paper currency and an additional driver towards a Cashless Society is transparency. I.E. making harder to some people to undertake on widespread anonymous and

untraceable transactions over and over again. The author (Rogoff, 2016) also argues that corruption prevents a society to become cashlessness.

We understand that cash can fuel the hidden economy and permit large-scale tax evasion, although the actual amount is, by definition, unknown (Achord, Chan, Nardani, & Rochemont, 2017). Also, it is difficult to track when it comes to tax collection or law enforcement (Bátiz-Lazo & Efthymiou, 2016). The anonymity of cash also makes it prone to criminal uses, which may explain why a lot of cash is issued in very high denominations not generally used by individuals or businesses (Broløs, 2016).

Primarily, the elimination of physical cash could earnestly impair criminal activity, particularly those connected with money laundering, corruption and drugs (Fabris, 2019). Those type of activities can be hardly carried out without cash (ibid). It is commonly believed that the use of cash enables privacy in transactions and can and does assist in the evasion of taxation (Achord, Chan, Nardani, & Rochemont, 2017).

Discussion around the substantial economic and social benefits that may result from transparency of a cash-free society and the end of physical cash, comprises the eradication of numerous of the most violent and serious crimes, significant cuts in taxes and better public services (Warwick, 2004). The reduction of the shadow economy will benefit the (digital) vaults of the countries by increasing public revenues, with the end result being the consolidation of their financial stability (Fabris, 2019). Most of the underground economy or black-market nowadays comprises undeclared transactions that could else be taxed (OECD, 2002). With the transition towards a cash-free society, such transactions would have to get in legal streams and be subject to taxation (Fabris, 2019).

2.2.6. Economic development & Financial Inclusion

It is not clear whether modern payment systems, which are developing at different speeds and with different outcomes in different countries, will help or exacerbate the problem of financial exclusion (OECD, 2002). However, presumably according to Kenneth Rogoff (2016), the more developed the countries are the less exposed is society to the risk of financial exclusion, when in the process of phasing out paper currency. It is crucial that unbanked and poor individuals have access to free standard debit accounts, and ideally also to basic smartphones (ibid). These costs are ideally borne directly

by the government or can also be imposed on banks that sooner or later will pass the costs on to paying clients (ibid).

Financial exclusion is identified as a problem for many people in both developed and developing countries (Achord, Chan, Nardani, & Rochemont, 2017). This is especially true for elderly that have used cash in all their lives and are likely to continue doing it as long as they will make payments (Arvidsson, 2019). The more impoverished and older population still remain disproportionately dependent on cash (Fabris, 2019). Since their knowledge of the use of digital technologies is usually more limited, the main concern is how the majority of them would manage in a cashless society (ibid).

However, the opposite effect in terms of financial inclusion should not be excluded (Fabris, 2019). Specially, in some rural areas or remote parts of a country with very limited financial infrastructures, digital money could lead to an increase in financial inclusion (ibid). The case of M-PESA in Kenya is a perfect example of that (Akinola, 2012). Not so long ago, payment cards embodied an international strategy to increase financial inclusion by allowing ready access to financial markets for low income and remote populations located far from economic centres and banks (Bátiz-Lazo & Efthymiou, 2016). In developing countries, the security concern to ordinary people carrying cash has become an additional motivation to address the financial exclusion of the vast unbanked sector (Akinola, 2012).

2.3. Summary of Literature Review and Justification for the Study

As exhaustively discussed, several academics have focused on the different drivers that could lead to a Cashless Society, although no studies made have explained to what extent each driver affects the level of cashlessness of each country and how they presumably could interact with each other, which, leaves scope for future studies.

The predominant findings consider a future Cashless Society as the result of multiple complex interactions between different stakeholders. In particular, it can be highlighted that:

- A Cashless Society is still a concept but one that could surely become a phenomenon over the next decades, due to the level of growth in the number of digital payments registered, at least

for a selected group of countries such as - Sweden or Canada (Fabris, 2019; Garcia-Swartz et al., 2006).

- The road to cashlessness is fuelled by technological innovations combined with continuous economic and social changes (Arvidsson, 2019; OECD, 2002).
- Due to the complexity in the transformation to a Cashless Society, one cannot attribute the change to solely one factor, rather during the transformation process, there are multiple stake-holders such as - banks, governments, individuals, merchants, fintech, among others (Arvidsson, 2019; Batiz-Lazo et al., 2016)
- Based on this literature review, the main hypothetical drivers are: **degree of digitalization** (Akinola, 2012; Fabris, 2019; Batiz-Lazo et al., 2016; Hedman & Henningsson, 2015); **the level of trust in digital technologies and privacy concerns** (Achord et al., 2017; Smithin, 2000); **legal framework** in terms of compliance and cybersecurity protection (Akinola, 2012; OECD, 2002; Rogoff, 2016; Warwick, 2004); **development of the banking industry** (Arvidsson, 2019; Rouse & Verhoef, 2016); **level of transparency and corruption** (Achord et al., 2017; Baubeau, 2016; Rogoff, 2016); and **economic development and financial inclusion** (Fabris, 2019; OECD, 2002; Rogoff, 2016).

These indispensable insights will serve as a basis for the implications analysed in the research design section (Methodology part) that will give rise to the research hypotheses tested in this thesis.

Our summary of the prevailing literature review reveals there is a need for a quantitative research model. Our quantitative research model should contribute to explaining the interactions among the hypothetical drivers and the resulting outcomes of a move towards a Cashless Society.

3. Methodology

The aim of this chapter is to provide a detailed description of all aspects of the design and procedure of this study, whose end purpose is to statistically validate or reject the abovementioned hypothesis as an attempt to find an answer to the initial research questions. After explaining the scientific assumptions that guide the selection of the chosen methods, this section will examine the research design so that the reader can judge the extent to which this paper would adequately answer the research questions. Then, the variables chosen to represent each of the hypothesised drivers as well as the subsequent outcomes will be described, along with the resulting research sample. Finally, this chapter will focus on justifying the statistical methods or procedures used to come up with the resulting findings.

3.1. Rationale for research question

As pointed out in the literature review, the topic addressed in this research shows a high degree of complexity and is affected by an uncountable number of factors that interact and influence each other. Assuming, therefore, that knowledge in social sciences is conjectural -no absolute truth can never be found-, this research takes a **postpositivist approach** (Creswell, 2014). This means that it intends to simplify the understanding of the cashless phenomenon by identifying those main common drivers across countries and statistically determine whether they affect the level of cashless transactions and its growth or not and, if so, whether they behave as positive drivers or, by contrast, as inhibitors.

First of all, this approach implies the recognition of potential biases, which need to be successfully minimised along the process: from the formulation of the initial hypothesis to the choice of certain statistical methods without underestimating the data collection procedure. As a consequence, this chapter pays special attention to justify the measures taken at every stage to prevent the authors from such biases and give validity and credibility to the final results.

Furthermore, the research needs to be designed in a way that allows further research to validate the final findings by either replicating the study with updated information or using alternative data sources to verify the concluding correlations between the hypothesised drivers and the outcomes.

Finally, it is important to emphasize that the initial hypothesis are based on an exhaustive review of the current literature. The postpositivist approach of this paper leaves room for additional drivers or inhibitors that could complement or update the current set of drivers as long as they are consistent with this research design and they are statistically valid. These new drivers could either emerge due to the impossibility to identify them with the current literature or changes in social, economic or technological terms, among others.

3.2. Research design

The design of this research seeks to give an answer to the research questions by statistically validating the initial hypothesis with a deductive reasoning. This implies that the study needs to be essentially quantitative, taking the assumptions of the postpositivist paradigm as a starting point.

The first step after having defined 6 hypotheses concerning the main potential drivers towards a cashless society is to assign numeric values not only to the drivers, but also to the outcomes. This process needs to take into consideration the following requirements:

- **Values** need to be given **on a country level** and have a **global vision across continents**. In other words, the study will not focus on certain cities or regions since it intends to give a global perspective to the results.
- Variables chosen need to come from **credible and reliable sources** that guarantee an objective view on the corresponding topics. To the possible extent, the data used will come from international public institutions such as the World Bank or United Nations. As a second option, this research is going to recur to global and reputed organizations such as the World Economic Forum or Transparency International. Lastly, original datasets produced by academic publications will be used after validating their relevance.
- Variables need to be extracted from studies or databases that are expected to be **published or updated on a regular basis** in order to give continuity to the analysis of this study. This requirement already implies that **secondary data** is a basis of this research.
- Variables assigned to each driver need to **reflect a static and dynamic view**. If, due to the novelty of the subject, data is insufficient to give both views, variables need to at least provide a static vision and give room for the dynamic perspective based on the existing and future publications.

- Data collected needs to be **consistent timewise**. The attempt to provide the most recent data should not jeopardize the point-in-time consistency.

Once the data is collected following the abovementioned requirements, the resulting research sample needs to be arranged in order to homogenise variables and optimise the number of countries studied. This process leads to several research samples that try to satisfy different needs, ranging from the maximization of variables studied to the maximization of the number of countries included in the sample.

Although the different resulting research samples can be used to enrich the analysis of this thesis, only the research sample with the highest number of variables and lowest degree of uncertainties regarding imputation or variable omission biases is used for the core conclusions of this paper. However, the alternative samples can be helpful to support the resulting findings. Additionally, further tools such as Clustering methods or simple mappings representations can also help to provide a better understanding of the research sample.

Unless all the driver variables are independent between each other, which can be checked by calculating the Pearson's Correlation Coefficients, the research sample previously chosen needs to be scaled and dimensionally reduced in order to simply the subsequent analysis. Then, the most relevant Principal Components or Dimensions -with eigenvalues greater than 1 (Kaiser, 1960)- can be categorized based on the main variables forming each PC. This implies that variables corresponding to different drivers could be included in the same PC. The categorized dimensions need, then, to be analysed by using multivariate regression methods (GLM) that sheds light on the initial research questions.

In order to answer the first research question ("**What are the main drivers towards a Cashless Society?")** the regression methods used aim to validate or reject the initial six hypotheses by statistically verifying the correlation between the dimensions (which contain information about at least one driver) and the outcome variables. Additionally, the direction of these correlations as well as the significant levels intend to answer the second research question ("**To what extent do they affect the cashlessness of each country?"**). Considering the potential size of the samples, this paper sets a minimum p-value of 0,05 to give statistical validity to each hypothesis (Royall, 1986). Moreover, the

research design is restricted to linear relationships not only during the regression analysis but also in the previous phases of research sample analysis and dimension reduction.

Finally, it is necessary to mention that Microsoft Excel has been used for the data collection and aggregation process considering that most of the databases were available in xls. or csv. formats. As for all the statistical methods, R -a free software environment for statistical computing and graphics- has been used.

3.3. Data collection

This section provides detailed information about the data collection process and the resulting variables, which needs to fulfil the constraints described in the research design. The division of this section into seven subsections corresponding to the variables of the outcome and the six drivers intends to give the reader a good understanding of the decision-making process that the authors of this paper undertook. Finally, an additional subsection gives an overview on all the variables described.

3.3.1. Variables – Outcome: level of cashlessness

In the previous chapter, the literature review has extensively explained what can be understood as a cashless society and has also introduced the term cashlessness. Converting the level of cashlessness of each country and its evolution over the last years into continuous variables represented a major challenge since these two dependent variables are the basis of the resulting findings.

Conceiving this phenomenon as a transitional process from a society that uniquely uses cash to a society that no longer uses it as a payment or exchange method may represent an oversimplification of a social change. However, this conception might be very helpful for the purpose of this study since the static outcome (level of cashlessness) could be represented by the percentage of non-cash transactions out of the total of transactions and the dynamic outcome could be represented as a growth of this percentage during the last years.

After a deep research, the authors concluded that this initial thought had to be discarded for a number of reasons:

- The data available was only provided for certain regions or cities or, in the best scenario, for individual countries.
- If the data had been collected country by country, the criteria of each source might have significantly differed from each other.
- The calculation of this percentage is always based on an initial estimation of the number of cash transaction, which in many countries is not feasible to provide with a high degree of certainty.

As a consequence, an alternative indicator had to be identified. In this case, the Global Payment Systems Survey (GPSS) from the World Bank might be helpful. This survey, initially launched in 2007, provides a regular payment systems overview of both advanced and emerging countries, by combining qualitative and quantitative measures. Among these measures, the survey included the volume of retail digital transactions in unitary and USD terms by country and transaction type (cash transactions are not included in this survey).

Although the fifth and latest survey was launched in 2018, the results were not expected to be available until, at least, late 2019. The last survey (World Bank, 2016), published in 2016, had information gathered on a yearly basis from 2010 to 2015. This issue represented, on the one hand, one of the trade-offs already anticipated during the research design (time consistency vs. recency). On the other hand, it also was giving the opportunity to other researchers to validate this study in the short term. After considering these two factors as well as the fact that the dataset allowed both a static and dynamic overview of the level of cashlessness and the notoriety and credibility of the institution supporting this initiative, the authors decided to opt for this source.

First of all, the data had to be cleared in order to make sure that certain transactions did not appear duplicate. For instance, in certain countries, there was information regarding card transactions and, additionally, details by card type. A similar case occurred with e-money, which in some cases was also given by channel. Based on the data available and cleared, the authors defined non-cash transactions as the sum of the following retail operations:

- Cheques
- Card payments
- Credit transfers

- Direct debits
- E-money transactions (including mobile and internet payments)

After clearing the data, the following step was to decide how the variable should be displayed. In order to reduce currency value and country size effects, the authors opted for using the sum of transactions in unitary terms and divide them by the country population on its corresponding year. By doing that, the first outcome variable would be obtained: non-cash transactions per capita, showing the level of cashlessness of a country in 2015 (static view), expecting the USA, Norway and South Africa, in which the last year available was 2014. This variable was called C1 in the dataset.

As for the second outcome variable (called C2), the compounded annual growth of non-cash transactions per capita between 2010 and 2015 by country was calculated (dynamic view), with the same exceptions explained above.

The resulting dataset consisted of 85 countries from all the continents ranging from developed economies and cashless pioneers such as Singapore or South Korea, to emerging economies with a low level of cashless transactions but a high growth such as Nigeria, Zimbabwe or Sri Lanka.

3.3.2. Variables – Driver 1: Degree of digitalization

Moving to variables of the drivers, the first diver to be codified is the degree of digitalization. In this case, the authors opted for the Digital Evolution Index, which, with the purpose of understanding how different countries are making the transition from a physical past to a digital future, it offers a simple means to measure at which stage of the transition each country is and how quickly countries are digitalizing. This Index, included in the Digital Planet report, takes 2008 as the starting point of the research and has been published in 2014 and 2017, respectively, as is expected to release regular future updates. Digital Planet is an initiative of The Fletcher School at Tufts University, in partnership with Mastercard, whose reports have had a relevant impact through the publication in the Harvard Business Review.

Coming back to the Digital Evolution Index, this data-driven holistic Index evaluates the progress of the digital economy across 60 countries, combining more than 100 different indicators across four key drivers:

- Supply Conditions, which includes transactions infrastructure indicators regarding the access to financial institutions and electronic payment options.
- Demand Conditions, which, among a wide range of indicators refer to the degree of financial inclusion and the use of digital money as well as the gender digital divide and the ability and willingness to spend.
- Institutional Environment, whose indicators range from the legal framework and the institutional effectiveness regarding the digital ecosystem.
- Innovation and Change, which refers, for instance, to financing options and opportunities.

Considering the presence of indicators regarding digitalization on each driver, this Index can be considered a good reflection of the degree of digitalization of each country. Additionally, as mentioned before, the publication of this Index not only shows a view of a certain year, but also the speed that countries are improving their level of digitalization. This duality cleary helps the purpose of this paper.

Therefore, the relevance and crebidility of this Index, the number and variety of countries covered as well as the possibility to not only give a static but also a dynamic view on this topic made the authors opt for this source.

At this stage, it is important to mention that Digital Planet changed the methodology of this Index in 2015. In order to avoid justifying the suitibility of each methodology, this driver includes four variables. The first two variables (**D1** and **D2**) show the degree of digitalization in 2015 and its 2008-2015 evolution, respectively, based on the new methodology (Chakravorti & Chaturvedi, 2017). As the other two variables (**DEI1** and **DEI2**), they show the digitalization score in 2013 and its 2008-2013 evolution based on the previous methodology (Chakravorti, Tunnard, & Chaturvedi, 2014). Although all these variables were included the the dataset, this study gives more validity to D1 and D2, not due to the methodology used, but to the time consistency with the other variables.

3.3.3. Variables – Driver 2: Digital Trust & Privacy Concerns

Although trust is an attribute very difficult to score due to its subjectivity, publications from OECD or The World Value Survey intend to assign cross-country scores to different types of trust such as the self-reported trust or the trust in public institutions (Ortiz-Ospina & Roser, 2019). However, the

way these types of trust are defined differs from the nature of this driver. Alternatively, Digital Planet (Chakravorti & Chaturvedi, 2017) built a new Digital Trust model based on four dimensions that answer different questions:

- Digital Environment: "How robust are privacy, security and accountability measures?"
- Digital User Experience: "How do users experience the digital trust environment?"
- Attitudes: "How do users feel about the digital trust environment?"
- Behaviour: "How do users react to their digital experiences and the environment?"

From these four dimensions, when measuring digital trust, there is an important distinction between behaviour and attitude since, as the report concludes, what users profess towards digital technologies (attitude) is not correlated with how the users actually behave, known as digital trust-in-action. Based on this, this study used behaviour score as the primary variable for this driver (**T1**). Additionally, a new variable (named **T2**) was created based on the gap between revealed preferences (attitude scores) and actual actions (behaviour scores) as an attempt to see if there was any correlation between this gap -which can be either positive or negative- and the growth in cashlessness.

These two variables, with data available for 42 countries and gathered in 2015, could give a good understanding of Digital Trust and Privacy Concerns, which as could be seen in the abovementioned report, were included in the design of the Digital Planet Trust model.

3.3.4. Variables – Driver 3: Legal Framework

The legal framework of each country can be evaluated from different perspectives. In this case, two of the most relevant legal aspects affecting digital payments are Compliance regulation and Cybersecurity protection. Additionally, the cross-country approach of this paper not only required credible evaluations, but also unified criteria. In order to fulfil these requirements, this study used two different sources that could enrich this driver.

As for compliance, FATF (Financial Action Task Force) and its regional bodies regularly update the Consolidated Assessment Ratings (FATF, 2019), which consist of peer reviews conduction on an ongoing basis to assess how effectively their respective member' AML/CFT measures work in practice, and how well they have implemented the technical requirements of the FATC Recommendations. This assessment brings an overview of 75 countries on both effectiveness and technical

compliance against the 2012 FATF Recommendations and in accordance with the 2013 FATC Methodology.

As mentioned before, the consolidated table provides information regarding two indicators. The first indicator (technical compliance) reflects the extent to which a country has implemented the technical requirements of the FATF Recommendations. This indicator is composed of 40 recommendations rated with C (Compliant), LC (Largely compliant), PC (Partially compliant) or NC (Non-compliant). In order to give a good understanding of the indicator in numeric terms, each rating was given a value ranging from 3 (Compliant) to 0 (Non-compliant). The sum of the 40 ratings was defined as Technical Compliance (**R1**), being 120 the highest level of technical compliance. For each country, the rating used was the one reported with the closest date to 2015 when this year was not available.

The second indicator (Effectiveness) reflects the extent to which a country's measures are effective. In this case, this assessment is conducted on the basis of 11 immediate outcomes, which represent key goals that an effective AML/CFT system should achieve. The rating of these immediate outcomes was given in a similar way: HE (High level of effectiveness), SE (Substantial level of effectiveness), ME (Moderate level of effectiveness) and LE (Low level of effectiveness). Therefore, the same procedure was followed by giving a certain score to each outcome, which ranged from 3 (High level of effectiveness) to 0 (Low level of effectiveness). The sum of the 11 outcomes was defined as Compliance Measures Effectiveness (**R2**), being 33 the highest level of effectiveness. As done in R1, the rating used for each country was the one reported with the closest date to 2015 when this year was not available.

With regard to Cybersecurity Protection, ITU and ABI Research jointly collaborated in 2013 in order to elaborate and produce the first publication of the Global Cybersecurity Index that was published in 2015 (ABI Research & ITU-UN, 2015). This Index has had further iterations published in 2017 and 2018 by ITU, the United Nations specialised agency for information and communication technologies.

This Index (GCI) is aimed at measuring the cybersecurity development capabilities of each country along five pillars:

- Legal Measures: scored based on the existence of legal institutions and frameworks dealing with cybersecurity and cybercrime.

- Technical & Procedural Measures: scored based on the existence of technical institutions and frameworks dealing with cybersecurity.
- Organizational Measures: scored based on the existence of policy coordination institutions and strategies for cybersecurity development at the national level.
- Capacity Building: scored based on the existence of research and development, education and training programmes; certified professionals and public sector agencies fostering capacity building.
- International Cooperation: scored based on the existence of partnerships, cooperative frameworks and information sharing networks.

The aggregation of these five pillars results in the final GCI score. The GCI report published in 2015, was based on data collected in 2014 across 182 countries and also included the GCI commitment score for 2017. Comparing the commitment scores for 2017 and the actual scores in 2017 (ITU, 2017) would show how reliable the commitment score was.

The collection of the GCI scores resulted in the following variables:

- **CY1:** Global Cybersecurity Index – 2015 Score
- **CY2-CY6:** GCI Parameters → Cybersecurity legal, technical, organizational, capacity building and cooperation measures – 2015 Score
- **CY7:** 2017 Cybersecurity commitment – 2015 Score
- **CY8:** GCI – 2017 Score

Due to the fact that all these indicators (R and CY variables) have been published from 2015, this study cannot include the dynamic component of this driver. However, this component can clearly be added in future publications that use the same or a similar research design since there are already evidences of data corresponding to the subsequent years.

3.3.5. Variables – Driver 4: Maturity of the banking industry

The development of the banking industry can be measured with a wide range of criteria. In this case, this driver aims to show how developed or mature the banking industry of a certain country is qualitative terms. In other words, the indicators show not directly depend on the size of the country.

The World Economic Forum publishes on a yearly basis The Global Competitiveness Index. The 8th pillar that composes the GCI is the Financial Market Development, defined in terms of (A) efficiency and (B) trustworthiness and confidence, which suits the description of this driver.

Based on the 2015-16 report (World Economic Forum, 2015), the first variable chosen was the Financial Markets Development Index (**B1**) with the 2015 Score. Since the report also provides historic data from 2005, the second variable chosen (**B2**) seek to reflect the evolution by calculating the compounded annual growth of the index between 2010 and 2015.

Whereas the first variable included scores for 140 countries, B2 only had information for 131 countries due to the lack of data regarding 2010 Financial Markets Development Index.

3.3.6. Variables – Driver 5: Transparency & Corruption

As seen in the Literature Review, Transparency and Corruption are two concepts that are used interchangeably by some of the sources earlier used. Since validating this correlation was out of the scope of this paper, the relationship presumed by this report is based on already existing empirical results, which conclude that transparency contributes to reduce corruption (Chen & Ganapati, 2018; Lindstedt & Naurin, 2006). In other words, both transparency and lack of corruption are not directly correlated, but transparency can be considered as an important trigger towards a lower level of corruption.

Said that, this fifth driver, named Transparency & Corruption aims to show the level of (lack of) corruption by country. This indicator is well-represented by the Corruption Perception Index from Transparency International, that conceives this index as an annual overview of the relative degree of corruption across countries all over the world on a scale from 0 (perceived to be highly corrupted) to 100 (perceived to be very clean). Therefore, by scoring the outcome (Corruption Level), the level of Transparency is to some extent also reflected.

In this case, CPI Scores published in 2016 (Transparency International, 2016), with data from 2015, included information regarding 168 countries. This Indicator was used as the first variable of this driver (**CP1**).

As for the second variable (**CP2**), which was intended to show the evolution of this last years, the compounded annual growth between 2010 and 2015 Scores was calculated, resulting in a dataset of

163 countries. Positive values in this second variable would mean an improvement in the degree of perceived corruption.

3.3.7. Variables – Driver 6: Economic development & Financial inclusion

This driver intends to provide a good understanding of the development of a country and, in particular, the level of financial inclusion. In order to enrich that final dataset, both variables regarding Economic Development and Financial Inclusion will be added.

With regard to Economic Development, as already been discussed in previous drivers, the variables should not be directly influenced by the size of the country. Therefore, relative indicators should be used. An important indicator of economic performance to make cross-country comparisons of average living standards and economic wellbeing is the GDP per capita. Based on data from the World Bank´s World Development indicators (World Bank, 2015), the first two variables chosen were the 2015 GDP per capita in current US$ (**GC1**) and the 2010-15 GDP per capita compounded annual growth (**GC2**).

These two variables used to score the Economic Development, were complemented with the 2015 Global Competitiveness Index (World Economic Forum, 2015). This report assesses the competitiveness landscape across countries with the aim of "providing insight into the drivers of their productivity and prosperity".

The final Index is the result of the aggregation of 12 pillars: Institutions, Infrastructure, Macroeconomic development, Health and primary education, Higher education and training, Goods market efficiency, Labour market efficiency, Financial market development, Technological readiness, Market size, Business sophistication & Innovation. Considering this wide range of factor, the use of this Index clearly enriches the oversimplification of the previous indicator (GDP per capita).

Therefore, two new variables were added. The first variable (**E1**) showed the GCI Score in 2015, whereas the second variable (**E2**) was the calculations of the GCI compounded annual growth between 2010 and 2015. The resulting dataset was an overview of 130 different countries.

As for the second component of this driver, Financial Inclusion, the Global Findex database was of great help. In 2011 the World Bank launched this initiative with funding from the Gates Foundation,

which is used to track progress toward the World Bank goal of Universal Financial Access by 2020. The Global Findex report (Demirgüç-Kunt, Klapper, Singer, Ansar, & Hess, 2010). So far, this database includes data from 2011, 2014 and 2017.

Among all the indicators gathered in the Global Findex database, there are two widely used in the report to reach conclusions regarding the degree of financial inclusion. These two indicators are account ownership and digital payments received or sent.

The first indicator shows the percentage of respondents (over 15 years old) who report having an account at a bank or another type of financial institution or report personally using a mobile money service in the past 12 months. From this indicator, two variables were created (**F1** and **F2**) representing the 2014 Account Ownership percentage and the 2011-14 Account Ownership compounded annual growth.

The second indicator shows the percentage of respondents (over 15 years old) who report "using mobile money, a debit or credit card, or a mobile phone to make a payment from an account, or report using the internet to pay bills or to buy something online, in the past 12 months. It also includes respondents who report paying bills, sending or receiving remittances, receiving payments for agricultural products, receiving government transfers, receiving wages, or receiving a public sector pension directly from or into a financial institution account or through a mobile money account in the past 12 months". Unfortunately, this indicator was added in 2014. As a consequence, only the static variable could be added (**F3**).

3.3.8. Summary of the variables

Based on the previous subsections, Table 1 gives an overview of the variables chosen classified by Category (Outcome or Driver) and component's nature (Static or Dynamic).

As a reminder, **Static variables** refer to the situation of each country in a specific year, e.g. C1 represents the number of cashless transactions per capita in 2015 by country; whereas **Dynamic variables** refer to the evolution or growth of each country during a specific period of time, e.g. C2 shows the compounded annual growth of cashless transactions per capita between 2010 and 2015 by country.

The purpose of the table below is to illustrate in a comprehensive and simplified way the datasets selected to represent each driver or outcome. This will be helpful for the reader to easily follow the analysis, discussion and conclusions provided in the subsequent sections.

Classification of Variables	**Static**	**Dynamic**
Outcome: Level of cashlessness	C1	C2
Driver 1: Degree of digitalization	D1, DEI1	D2, DEI2
Driver 2: Digital Trust & Privacy Concerns	T1	T2
Driver 3: Legal Framework	R1, R2, CY1-CY8	
Driver 4: Maturity of the Banking Industry	B1	B2
Driver 5: Transparency & Corruption	CP1	CP2
Driver 6: Economic development & Financial inclusion	GC1, E1, F1, F3	GC2, E2, F2

Table 1: Classification of Variables by Category and component's nature

3.4. Unified Dataset & Research samples

This section is going to describe the dataset that resulted from merging all the collected data as well as the research samples used for the analysis.

3.4.1. Unified Dataset

This initial dataset was the starting point for the arrangement of the research samples that will be descripted in the following subsections.

This dataset -named Unified Dataset- was the result of the aggregation of the **29 variables** (2 Outcome/Response variables and 27 Driver/Explanatory variables). Additionally, this dataset only includes those country that, at least, had information regarding the outcome variables (C1 & C2).

During the aggregation process, each country name was standardized across datasets and was assigned the corresponding ISO 3166-1 alpha-3 code (ISO, 2019) in order to avoid mistakes during the variable datasets merger.

As a result, **54 countries** -from all the world regions- were included in this dataset as displayed in the following map - see Appendix A for a detailed list of countries (A.1) as well as a sample of the variable values (A.2) and a brief statistical description of each variable (A.3):

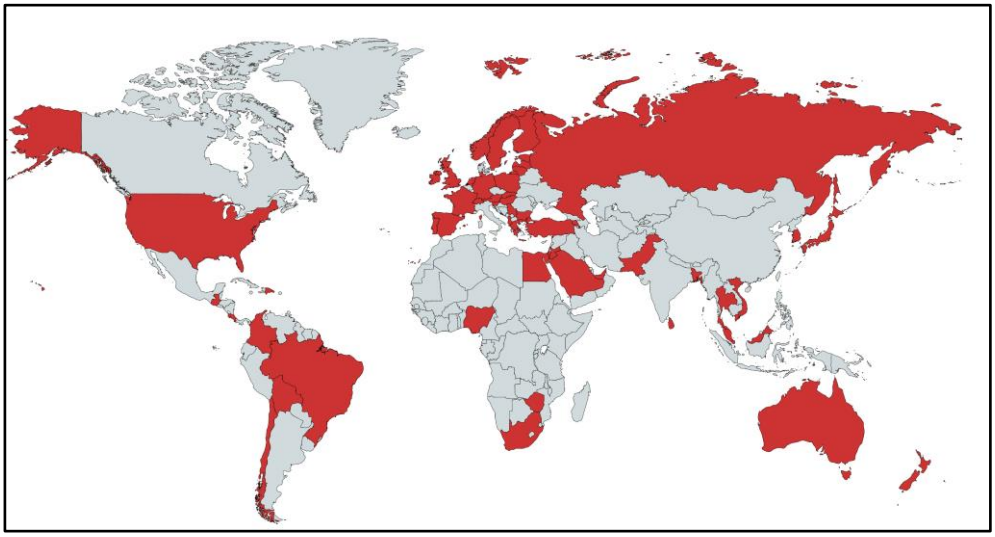


Figure 2: Map of countries included in the Unified Dataset – Authors' elaboration with MapChart.net

The Unified Dataset has a percentage of 9,2% of missing values accumulated in 9 of the explanatory variables as shown below:

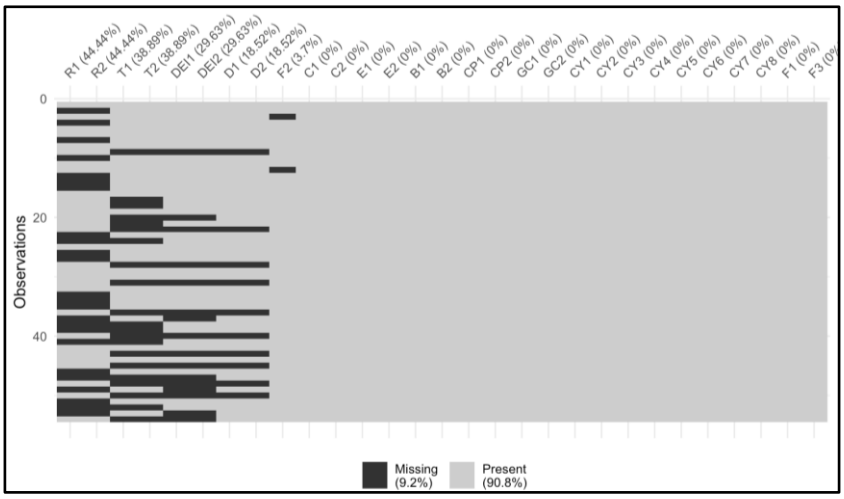


Figure 3: Representation of missing values in the Unified Dataset – Authors' elaboration with R.

In addition to a clear concentration of missing values in a few variables (6 variables have more than 25% of the values missing), when illustrating the distribution of missing values against C1 (see Appendix A.4), it can be perfectly seen that most of the data missing belongs to countries with a low level of cashless transactions.

3.4.2. Research sample with Complete Data

The first Research sample arranged is the result of filtering those countries without missing values and keeping all the variables included in the initial Unified dataset. As shown below, the final sample included **13 countries** across the continents, excepting Africa (see Appendix B.1 for details):

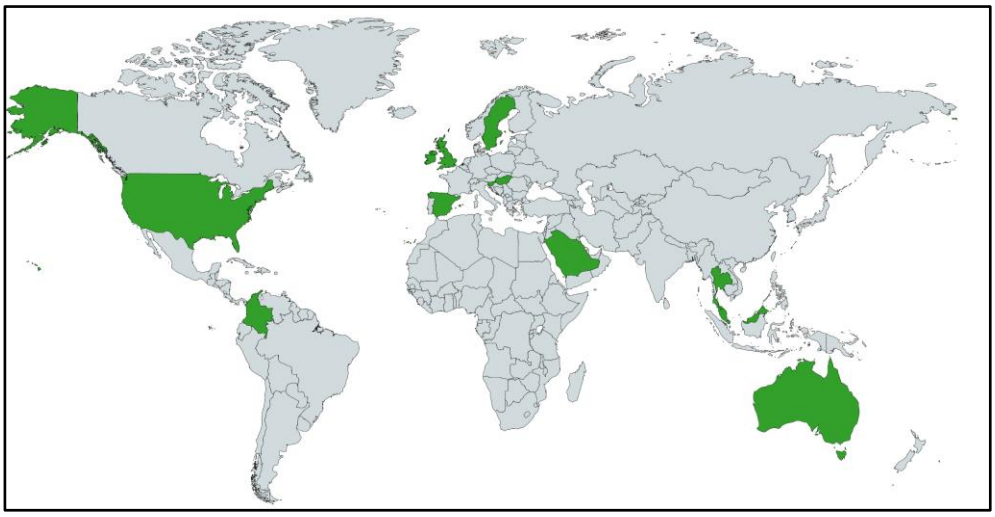


Figure 4: Map of countries included in the RS with Complete Data. – Authors' elaboration with MapChart.net

Although these 13 countries only represent 25% of the ones included in the initial Dataset, it is important to emphasize that they still are in different cashlessness stages and any variable has been omitted. Therefore, **this Research sample will be the primary source for the subsequent analysis in order to reach the main conclusions of this study**.

3.4.3. Research sample with imputed missing values

A solution to avoid reducing the sample size is to impute values to the missing data. After considering several imputation methods, the k-fold cross-validation method was chosen. The **K-fold Cross Validation Method** is based on dimensionality reduction methods in which the missing entries are imputed using the iterative PCA algorithms that take into account the similarities between the observations and the relationship between variables. It has proven to be very competitive in terms of quality of the prediction compared to other traditional methods in numerical experiments (Camacho & Ferrer, 2012).

Despite the potentiality of this method, the high concentration of missing values in underdeveloped countries with a low level of cashlessness, can lead to a significant imputation bias. For example, a

PC analysis made a posteriori shows that the variables with the highest number of missing values (T1, T2, R1, R2) are included in the same dimension, which is not consistent neither with literature review nor with a PC analysis of complete data.

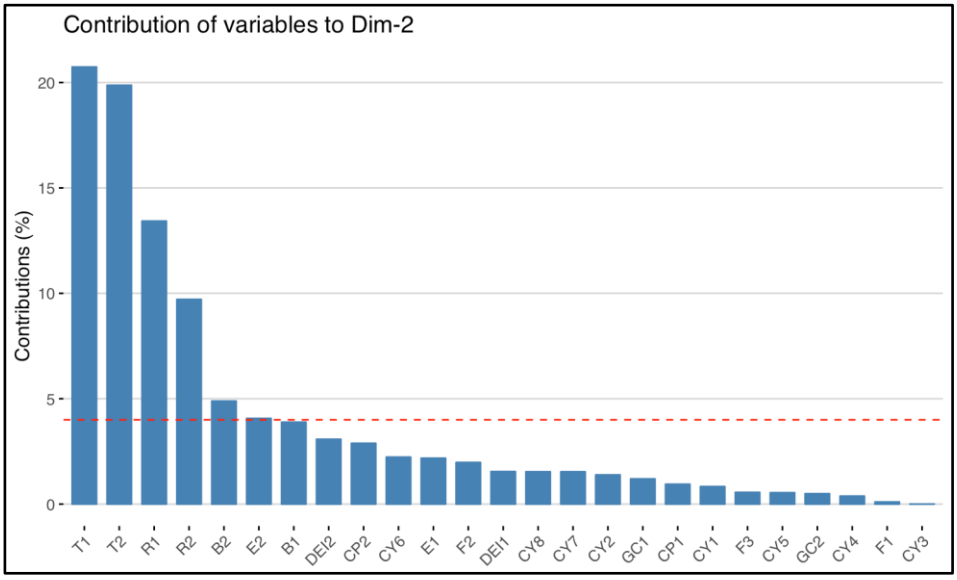


Figure 5: Second Principal Component (in terms of contribution %) obtained from a PC analysis of the Research sample with imputed missing values.

According to Liu & Gopalakrishnan (2017), the percentage of missing values and their distribution across variables are two points of consideration when answering the question whether to impute or not. Taking into account that these are two of the major concerns raised about imputing missing value with the dataset provided, it can be concluded that **this sample should not be used neither for the main nor for complementary analysis**.

3.4.4. Optimised Research sample

An alternative way to avoid the significant number of countries studied is to optimise the number of variables and countries by obtaining the complete dataset with the highest number of values possible and without having to impute data. The resulting dataset would omit those four variables with the highest percentage of missing values (T1, T2, R1 and R2) and would include **44 (out of 54) countries** across all continents (see Appendix B.2 for details):

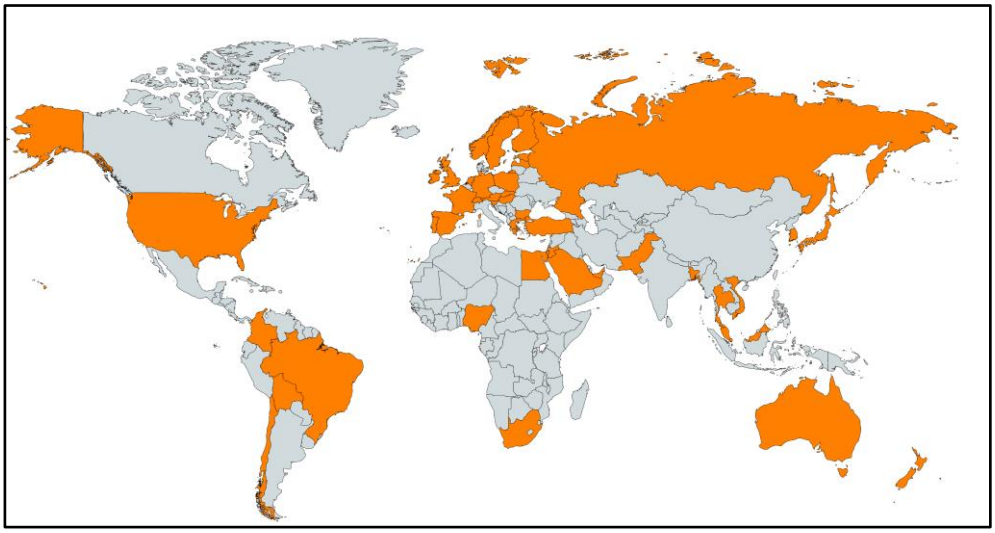


Figure 6: Map of countries included in the Optimised Research sample – Authors' elaboration with MapChart.net

The **omission of these four variables** that correspond to two drivers (2. Digital Trust & Privacy Concerns and 3. Legal Framework) either deletes or weakens the description of this drivers in the dataset. As a consequence, this sample cannot be used to reach conclusions about the whole model design. However, it can be used to provide further evidences about partial findings referring to the drivers whose variables have not been altered.

Therefore, **the Optimised Research sample will be used complement the analysis from the Research Sample of Complete data**, always considering the variable reduction constraint.

3.5. Data analysis methods

3.5.1. Preliminary data scalation

Prior to the application of many multivariate data analysis methods, data are often pre-processed. In this case, due to the differences in the value ranges, all the values from the research samples used have been scaled with the aim of improving predictive accuracy. In R, the default scale function standardizes the data (mean zero, unit variance).

In order to avoid losing information that could affect the resulting standardized values, this scalation process has been performed against the Unified dataset. As a result, all the Research samples had the same scaled values by country and variable.

3.5.2. Pearson's correlation between variables

Before moving to the core of this study, the calculation of Pearson's correlation coefficients can provide an initial understanding of the statistical relationship, or association, between the explanatory variables. These coefficients do not only give information about the magnitude of the association, or correlation, but also about the direction of the relationship.

The calculated coefficients can be used to stablish preliminary relationships between variables from different drivers and, additionally, validate the dimensions obtained in the PC analysis explained in the following section.

In order to calculate these correlation coefficients, the Unified dataset with scales values were used, but only pairwise complete observations were included.

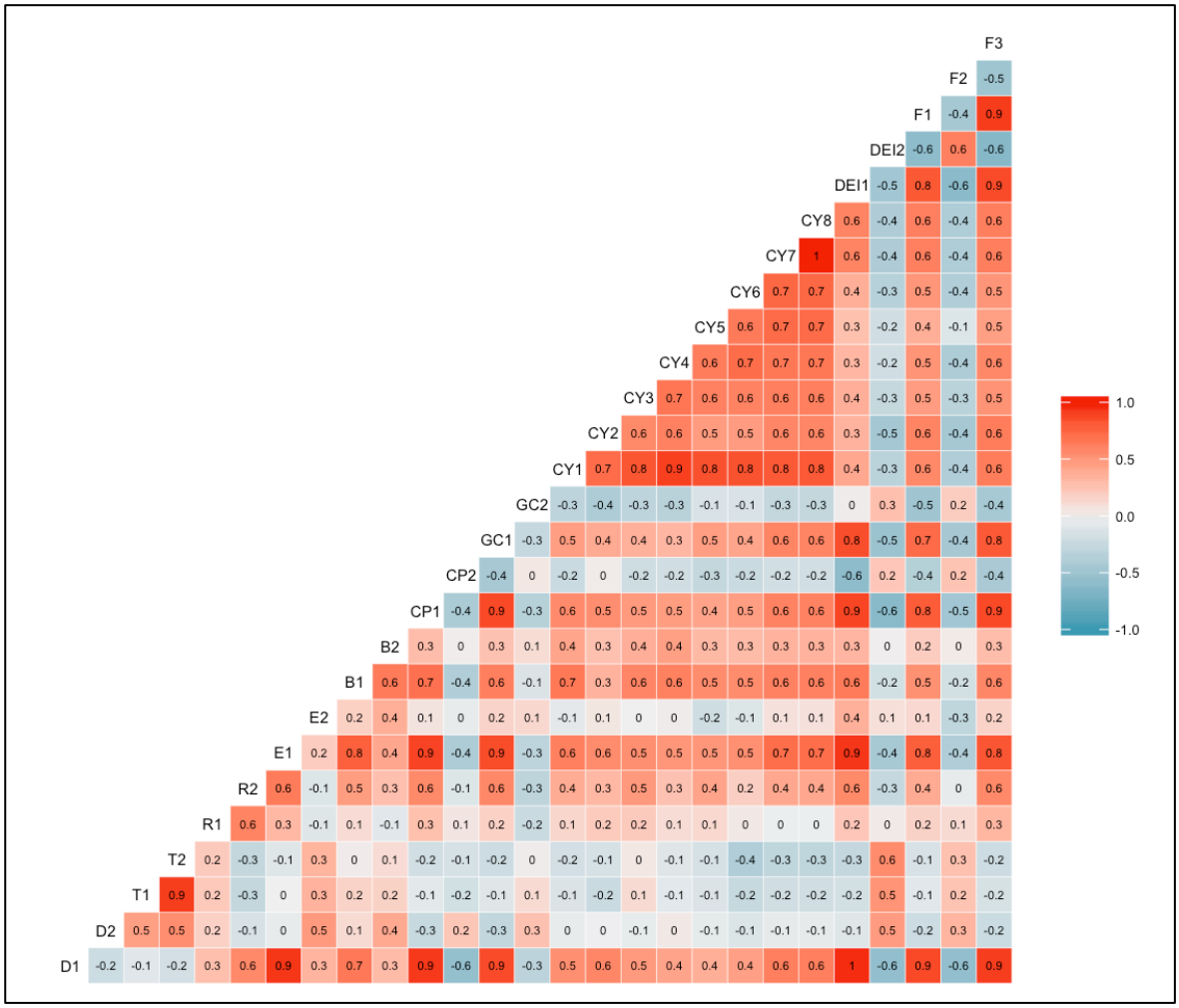


Figure 7: Pearson's correlation between variables

3.5.3. Principal Component Analysis

Considering the correlation between certain explanatory variables from different drivers and the large set of variables (Cramer, 2003), a dimensionality-reduction method such as Principal Component Analysis (PCA) has been used to reduce the set of explanatory variables to a smaller number of dimensions. By doing that, it will be easier to explain and visualize which is the nature of the chosen explanatory variables and how they can be grouped.

The PCA was run on the Research sample with Complete Data -the reasons for this choice have already been justified before- and it resulted in a plot of 10 dimensions out of which only 6 have Principal Components have eigenvalues higher than one (see Appendix C.1. for further details).

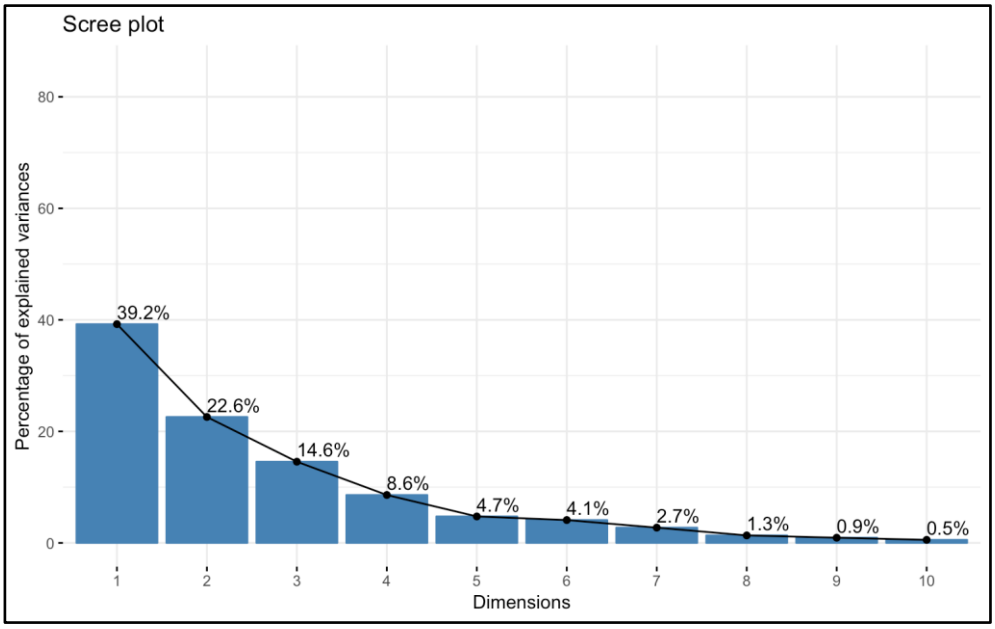


Figure 8: Scree plot showing the percentage of explained variances of each dimension. – Authors' elaboration with R

In order to provide more details about the contribution of the variables within each dimension, two different graphs were prepared for each relevant dimension (as shown in Appendix C.2) giving the contribution % of variables to each dimension and graphically illustrating the direction of this contribution, respectively. These graphs show help to categorize each dimension.

This PC Analysis on the Research Sample with Complete Data and the resulting dimension values are the basis for the following sections as it represents a simplification of the current sample without a significant loss of information.

As earlier mentioned, a similar PC Analysis was applied to the Research sample with imputed values to discard this sample as a valid dataset for the subsequent analysis. By contrast, the Optimised Research sample was also analysed using PCA methods (see details in Appendix C.3). In this case, the results were aimed at either supporting or rejecting the main conclusions reached with the results from the analysis of the Research Sample with Complete Data.

3.5.4. Cluster analysis on PCA results

Based on the PCA results, another way to illustrate the reasoning of the dimensions is to cluster the countries in order to find certain patterns. Therefore, in this paper, the single purpose of the hierarchical cluster analysis is to complement the core analysis.

Focusing on the hierarchical clustering, the optimal number of clusters will be determined by the Elbow methods unless it cannot be unambiguously identified (Kodinariya & Makwana, 2013).

This hierarchical cluster analysis has been employed in the Research Sample (RS) with Complete Data (see Appendix C.4 for further details) and the clustered countries have been graphically represented against the outcome variables (not included in the PC Analysis).

Furthermore, after considering that the first and second dimensions of PC analysis on the RS with Complete data and on the Optimised RS have very similar variable contributions, the hierarchical cluster analysis has also been employed in the Optimised RS in order to provide a graphical representation with more countries. However, as already stated before, it is important to bear in mind that the PCA results of this analysis and, thus, the clusters forms, do present a lack of information regarding two of the hypothesised drivers. As a consequence, the results of this second cluster analysis (see Appendix C.5 for details) can only be used as a complement of the core results.

3.5.5. Regression analysis

Coming back to the core analysis of this study, once all the variables have been aggregated in 6 dimensions (eigenvalues > 1), it is necessary to draw the influence of these drivers against the outcome or response variables.

In other words, the analysis method used needs to provide insights regarding the degree of influence (significance) of the drivers against each outcome variable (Research Question 1) and the direction of this influence in case there is a relevant significance (Research Question 2).

An optimal way to answer these questions is to employ a regression analysis in order to model the relationship between the dimensions of the RS with Complete Data and each response variable. In this case, two multiple linear regressions will be used.

Each multiple linear regression employed for each outcome variable will follow this formula:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_6 X_{i6} + \epsilon_i$$

for each observation $i = 1, ..., 13$

In the formula above 13 observations (corresponding to the 13 countries of the RS with Complete Data) of one dependent variable (either C1 or C2) and 6 independent variables (corresponding to the 6 dimensions) are considered. Thus, Y_i is the i^{th} observation of the outcome variable, X_{ij} is i^{th} observation of the j^{th} dimension, $j = 1, 2, ..., p$. The values β_j represent parameters to be estimated, and ε_i is the i^{th} independent identically distributed normal error.

The aim of this paper is not to get a restrictive formula for each dependent variable, but to see the direction of each β_j and its significance level (p-value) in order to determine those dimensions that have a significance influence on each outcome.

Based on these results (see Appendix C.6), those significant dimensions need to be disaggregated to identify to which drivers they correspond to and whether they belong to the dynamic or static component of each driver.

4. Main Findings

This section is going to expose those finding that resulted from applying the data analysis methods defined in the Methodology section. These findings will be extensively described in a comprehensive way following the logical order of discovery. So, at first the dimensions resulting from the PC Analysis will be presented along with their corresponding interpretation. Then, the results regarding the country cluster analysis will be explained. Finally, the outputs from the regression analysis will be highlighted and aggregated with the intention of simplifying the core findings of this thesis.

4.1. Dimension categorization

This subsection is going to analyse separately each of the 6 first dimensions with the aim of identifying those variables that contribute to a higher extent to the resulting dimension and thus giving a name or category to each dimension. Before examining each dimension, it is important to mention that each consecutive dimension represents a lower percentage of explained variances. Therefore, whereas for the first principal components a larger number of variables should be considered to assign a category, in the last cases only the primary variables with be considered.

Dim-1: Digital & Economic development

After calculating the Pearson's correlation coefficients (see Figure 7), it could be clearly seen that static components from all the drivers excepting Digital Trust & Privacy (Driver 2) were in many cases correlated with coefficients higher than 0,8. These one-to-one correlations augured well for the aggregation of most of these variables in one Dimension.

As shown in the contribution graph of the first dimension (Appendix C.2), the level of digitalization, (DEI1, D1), effectiveness of compliance regulation (R2) as well as cybersecurity protection (CY1, CY7, CY8), maturity of the banking industry (B1), transparency (CP1) and, finally, economic development & financial inclusion (E1, GC1 & F3) are well represented in this dimension.

This amalgam of variables makes the dimension categorization rather challenging. However, the authors of this paper finally opted for calling it Digital & Economic Development taking into account which are the variables with the highest contribution percentage.

To sum up, this first dimension fully represents the static component of Driver 1, Driver 4, Driver 5 and Driver 6 and partially represents the static component of Driver 3 as R1 is not significant. In this case, Driver 3 is represented in Dim-1 by R2 and CY variables, which refer to regulation effectiveness and cybersecurity protection

Dim-2: Digital & Transparency growth

Moving to the second dimension, there is a variable that clearly contributes more than the others: D2 – Digital Growth. This variable is mainly following by growth variables corresponding to other drivers, but with much lower contribution percentages.

Considering that the second variable corresponds to Corruption Perception (CP2) and that the level of digitalization has one of the highest correlations with the level of Transparency (or Lack of Corruption perceived), this feature has also been included in the final dimension description.

Within these other variables, there is also DEI2 – Digital Growth. However, its lower contribution could also be explained by the low correlation between these two Digital Growth variables. This differences probably come from the changes in the Methodology used to score these growths or a drastic growth change between 2013 and 2015 in some of the countries included in the sample.

Based on all these considerations, this second dimension has been called Digital & Transparency Growth and includes the dynamic component of Driver 1 and Driver 5.

Dim-3: Economic growth

In this case, the variable leading the dimension is the Growth in GDP per Capita (GC2), closely followed by E2 (Growth in the Global Competitiveness Index). These two dynamic variables, that belong to Driver 6, will be used to categorize this dimension naming it Economic Growth.

CY2, CY3 and CY4, variables that score Global Cybersecurity Index parameters will not be included in the definition of the dimension due to their inconsistency with the other variables. First of all, whereas the former variables represent a dynamic component, these latter variables show a static score. However, the main reason to discard these variables is the nature of these variables.

The variables that belong to GCI parameters (CY2-CY6) are the ones from the dataset with the weakest scale. Although they were chosen in order to enrich the dataset, the country reduction performed in the RS with Complete Data along with the variables scalation process have clearly affected these variables. As a consequence, they cannot be taken into consideration.

Dim-4: Compliance requirements

While R2, defined as the effectiveness of compliance measures had already appeared in Dim-1, R1 has a prominent contribution to this Dimension. As earlier described, R1 represents the level of technical compliance of each country.

Considering that the other variables have a much lower contribution percentage, R1 will be the only variable used to categorize this dimension as Compliance Requirements. This dimension is, therefore, only taking a static component of Driver 3.

Dim-5: Digital Trust

As earlier mentioned, these last two dimensions that represent less than 5% of the explained variances of the Sample will only use the first variable to define the dimension. In this particular case, since T1 describes the level of Digital Trust&Privacy Concerns, being the static component of Driver 2, this dimension will also be named Digital Trust.

Dim-6: Banking maturity growth

Finally, the last dimension is led by B2 – Growth in the Development of the Banking Industry. This dynamic component of Driver 4 is going to be used to define this last dimension as Banking maturity growth.

Overview of the dimensions

The table found below intends to give an overview of the relationship between the drivers and the dimensions based on those representative variables that gave name to each dimension. This table also shows whether the variables used for the dimension categorization belong to the static or dynamic component of each driver. In other words, those components marked with a “X” are the ones used to categorize each dimension.

Although this dimension categorization process leads to a loss of certain information, it clearly simplifies the understanding of each dimension.

Drivers \ ***Dimensions***		Dim-1: Digital & Economic Development	Dim-2: Digital & Transparency growth	Dim-3: Economic growth	Dim-4: Compliance requirements	Dim-5: Digital Trust	Dim-6: Banking maturity growth
Driver 1: Degree of digitalization	Static	**X**					
	Dynamic		**X**				
Driver 2: Digital Trust & Privacy Concerns	Static					**X**	
	Dynamic						
Driver 3: Legal Framework	Static	**X**			**X**		
	Dynamic						
Driver 4: Maturity of the Banking Industry	Static	**X**					
	Dynamic						**X**
Driver 5: Transparency & Corruption	Static	**X**					
	Dynamic		**X**				
Driver 6: Economic development & Financial inclusion	Static	**X**					
	Dynamic			**X**			
		Static	Dynamic	Dynamic	Static	Static	Dynamic

Table 2: Explanation of the categories given to each Dimension based on the classification of the representative variables.

4.2. Country clustering

After conducting the PC analysis, a cluster analysis has been performed in order to find certain patterns that could complement the core findings of this study. Initially, the Research Sample with Complete Data has been clustered based on the dimensions previously obtained. This cluster analysis has resulted into four groups.

Considering the primary drivers as well as their positions in the PCA-Biplots (Appendix C.2), these clusters could be grouped as follows:

- GROUP I: Countries with a low or medium Digital & Economic development (Dim-1), but with high Digital growths (Dim-2).
- GROUP II: Countries with a low or medium Digital & Economic development (Dim-1), but also with low or medium Digital growths (Dim-2) and low Economic growths (Dim-3).
- GROUP III: Countries with a medium Digital & Economic development (Dim-1), but with medium Digital growths (Dim-2) and high Economic growths (Dim-3).
- GROUP IV: Countries with a high Digital & Economic Development (Dim-1) and a high Digital growth (Dim-2).

When displaying these groups in a graph with C1-Cashless transactions per capita as X-Axis and C2-Compounded annual growth of cashless transactions per capita as Y-Axis, we can see that there is a fairly good match between the clusters and different stages of cashlessness:

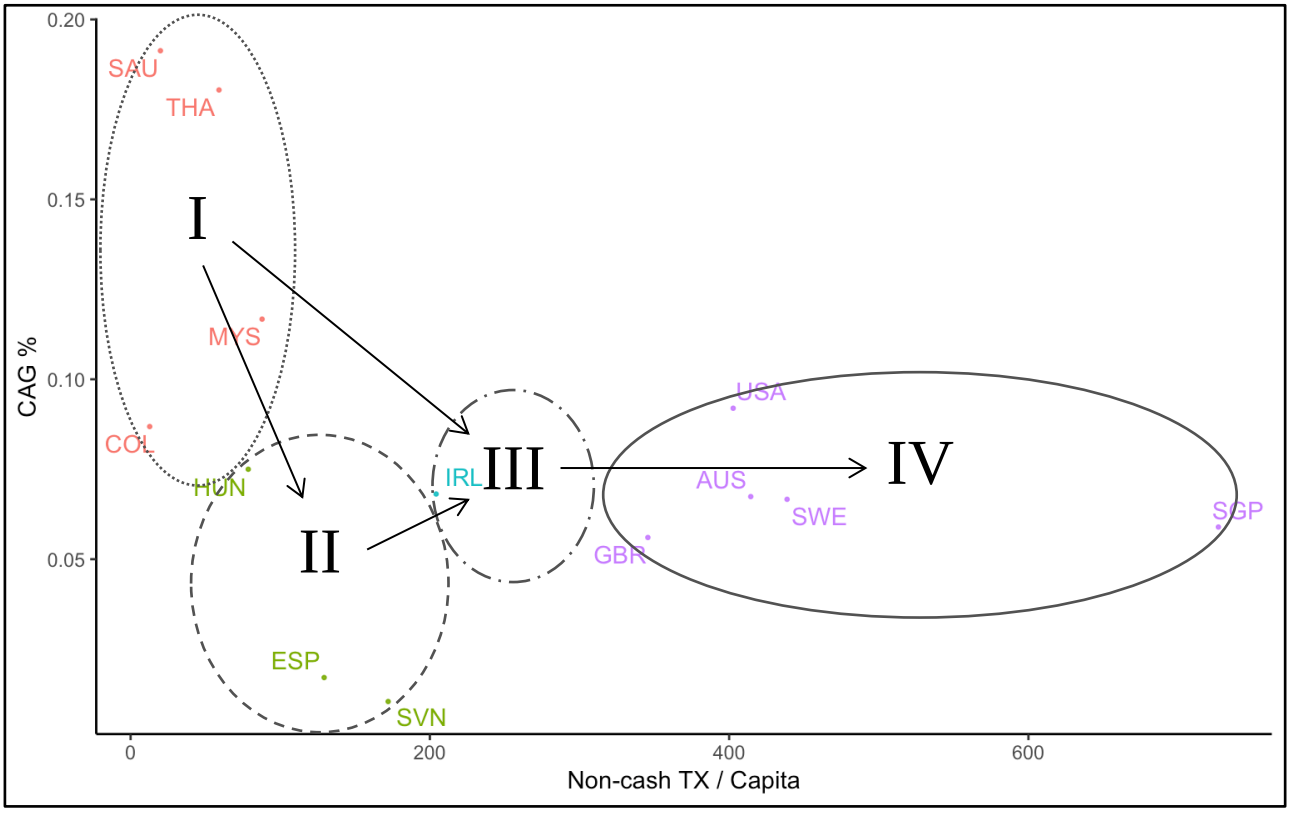


Figure 9: Graph displaying the clustered countries from RS with Complete Data by outcome variables. – Authors' elaboration with R

- GROUP I represents those countries in an early stage of cashlessness, but with high growths.
- GROUP II represents those counties with a moderate level of cashlessness, but with low or even negative growths (probably due to contextual reasons such as economic downturns).
- GROUP III represents countries with a moderate level of cashlessness and a steady growth.
- GROUP IV represents those countries in an advanced stage of cashlessness.

These groups are relative since the average level of cashlessness of the observed sample needs to be considered, both in absolute terms (C1) and relative terms (C2). At the same time, what can be concluded is that, even if some countries temporarily weaken their cashlessness positions due to contextual reasons, there is a trend to move towards a cashless society. However, there are not enough evidences to determine whether countries that belong to Group I are moving to become members of Group II or they would directly jump into Group III.

Said that, since the first and second dimensions of PC analysis of the Research Sample with Complete data and the Optimised Research Sample have very similar variables included, categories for the first two dimensions can be presumably shared between these two samples. Using again the response variables as axis, this is the display of clustered countries from the Optimised RS:

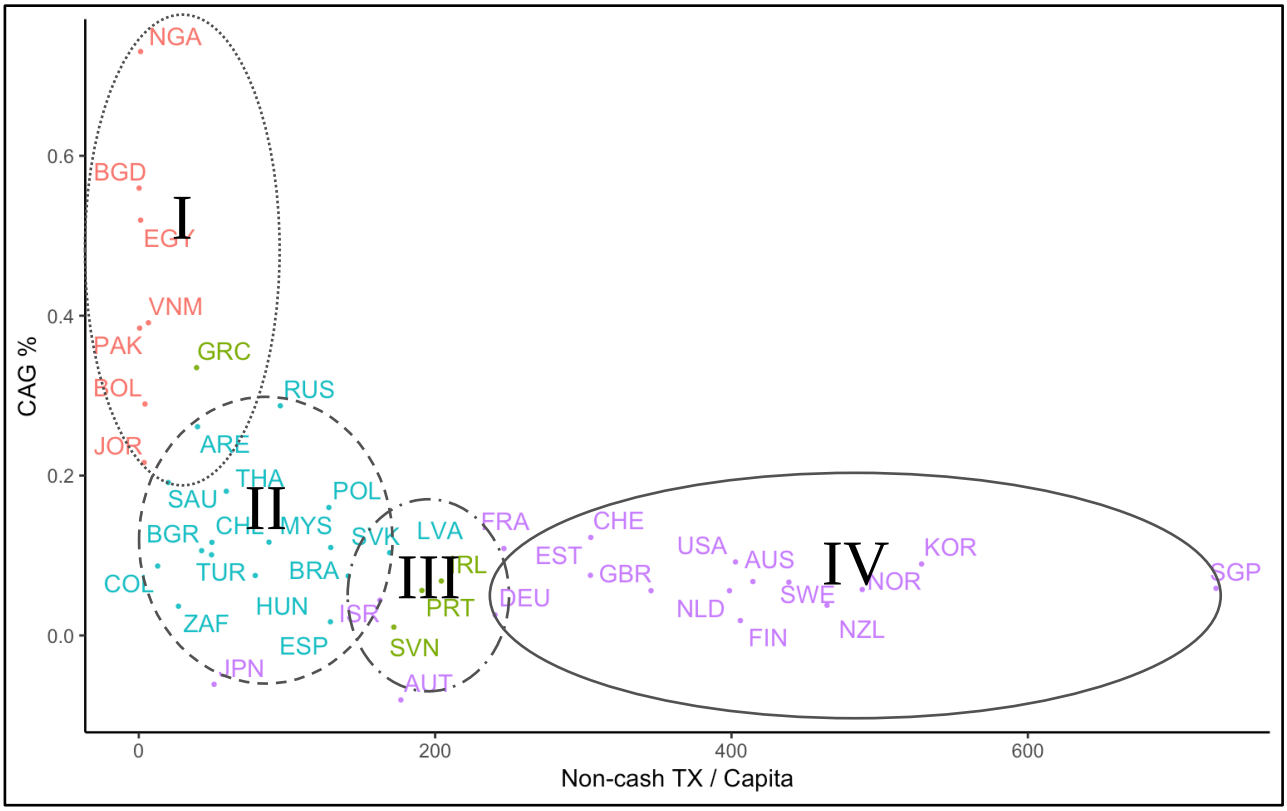


Figure 10: Graph displaying the clustered countries from Optimised RS by outcome variables – Authors' elab. with R

When overlapping the previous groups with the updated set of countries, there is not a clear match even though certain patterns are followed. Considering that this latter graph comes from the Optimised RS in with two dimensions were not included in order to amplify the list of countries, this could lead to the conclusion that more dimensions need to be considered when relating the features of a country with its cashlessness position. For this reason, the regression analysis is essential to reach substantial and significant conclusions.

4.3. Regression between dimensions and outcome variables

As stated before, this regression analysis is expected to give an answer to the initial research questions. This analysis has been conducted for each outcome variable (C1 and C2) in order to see the effects of the hypothesized drivers from both a static and dynamic perspective, respectively.

The results of this regression analysis (see Appendix C.6 for details) have been displayed in two separate graphs that simplify the understanding of the behaviour of each dimension against the corresponding outcome variables. First of all, each dimension is represented by a horizontal line whose length depends on its variance. In other words, the longer the line is, the more disperse the results are and the less chances to reach binding conclusions.

Additionally, the location of each line with respect to the dotted vertical line shows the type of effect between each driver and outcome. In some cases, a positive effect (right side of the dotted line) or a negative effect (left side) can be clearly observed. However, in other cases, the dimension line lies on both sides of the vertical line, which means that not all the countries show the same pattern.

To sum up, the shorter and the further apart from the dotted line the dimension is, the more meaningful the results are. In statistical terms, the level of significance is determined by the p-value as explained in the Methodology section (see Section 3.5.5 in the Methodology chapter).

The illustration of the results obtained from the regression analysis are presented hereunder by outcome or response variable.

C1 as response variable

The first regression analysis performed was against C1, being the static components of the outcome variables. The results of this analysis have been illustrated in the following graph:

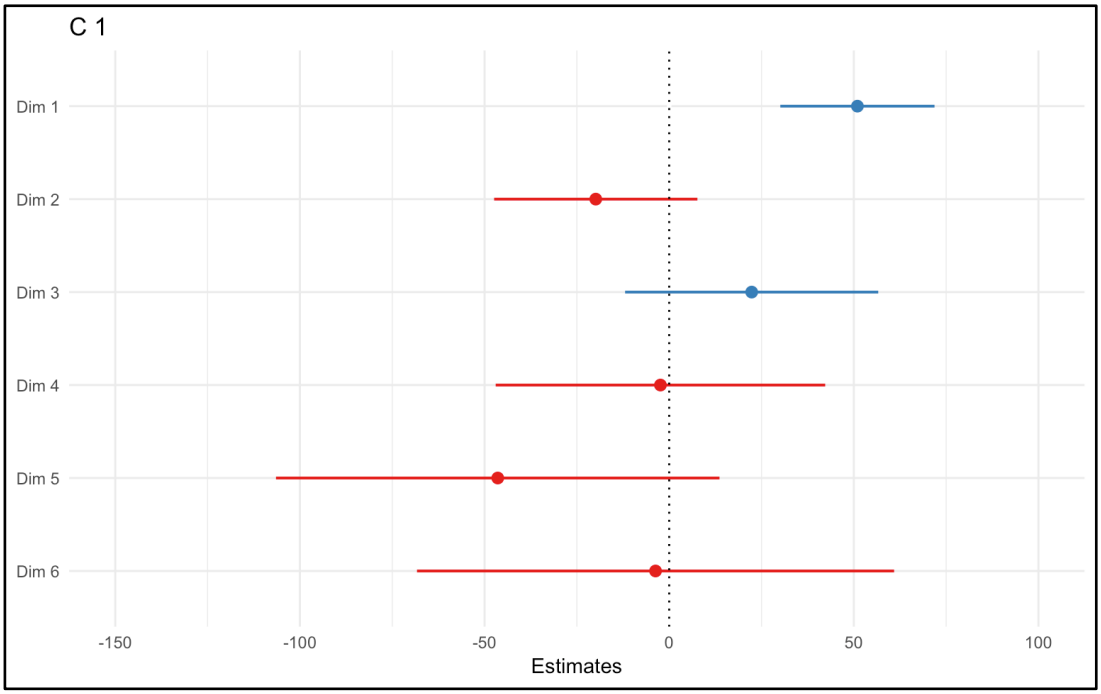


Figure 11: Illustration of the regression analysis against C1. – Authors' elaboration with R

This graph shows the degree of correlation of each dimension with C1. The shorter the line is, the lower the variance of the estimated parameter (β) is. However, the P-value of each dimension is going to be critical to determine which dimensions have a significant influence on the outcomes, as mentioned above.

In this case, the only Dimension that statistically has a direct positive influence on the number of cashless transaction per capita (C1) is Dimension 1, which includes static components of several drivers such as Degree of Digitalization or Economic Development.

The second lowest p-value comes from Dimension 5, which is probably determined by its high variance. Although the p-value is not statistically significant to reach binding conclusions about this dimension, this could explain the behaviour of certain countries towards a lower level of cashlessness despite its high Digital & Economic Development.

Therefore, based on the directions of the estimated parameters, C1 can be explained by the following formula: **+ Dim1 (– Dim5) → + C1** or, in other words:

+ Digital & Economic Development (– Digital Trust) → Cashless Transactions/Capita

When disaggregating each dimension by driver components, C1 can be explain as a result of:

+ D1s: Degree of digitalization
(– D2s: Digital trust)
+ D3s: Regulation effectiveness
+ D4s: Maturity of the banking industry
+ D5s: Transparency
+ D6s: Economic Development & Financial Inclusion

} **C1: Cashless Tx/Capita**

C2 as response variable

The second regression analysis performed was against C2, being the dynamic components of the outcome variables and representing the growth in cashless transactions per capita between 2010 and 2015. The results of this analysis have been illustrated in the following graph:

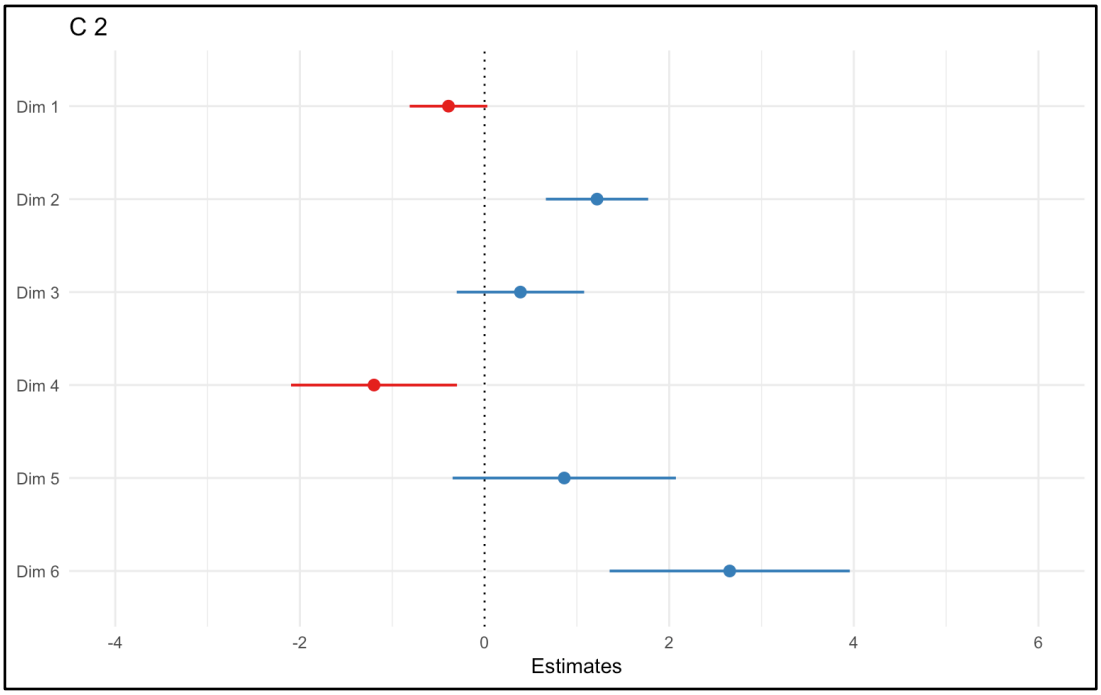


Figure 72: Illustration of the regression analysis against C2. – Authors' elaboration with R

Following the same reasoning as C1, this graph -along with the regression results- determines that Dimension 2 and Dimension 6 have a positive influence on C2. By contrast, Dimension 4 has a negative effect against C2. All these results are statistically significant considering their p-values.

Therefore, based on the directions of the estimated parameters, C2 can be explained by the following formula: **+ Dim2 – Dim4 + Dim6 → + C2** or, in other words:

+ Digital & Transp. Growth – Compliance Reg. + Bank. Matur. Growth → Cashless Growth

When disaggregating each dimension by driver components, C1 can be explain as a result of:

+ $D1_D$: Degree of digitalization
– $D3_S$: Compliance regulation
+ $D4_D$: Maturity of the Banking Industry
+ $D5_D$: Transparency

} **C2: Cashless Growth**

Overview of the regression analysis

Considering all these findings, the initial drivers stated in the hypothesis have the following effects on each of the Cashlessness components defined along the thesis:

Drivers' Effects on Cashlessness components	Static (C1)	Dynamic (C2)
Driver 1: Degree of digitalization	+	+
Driver 2: Digital Trust & Privacy Concerns	(–)	
Driver 3: Legal Framework	+	–
Driver 4: Maturity of the Banking Industry	+	+
Driver 5: Transparency & Corruption	+	+
Driver 6: Economic development & Financial inclusion	+	

Table 3: Overview of the correlation between hypothesised drivers and the cashlessness components.

Based on the summary of Table 3, both components can be merged into a unique parameter that shows the effect of the drivers on the overall level of cashlessness. **The following table summarizes the core findings of this thesis:**

Drivers' Effects on the overall level of Cashlessness	**Level of cashlessness**
Driver 1: Degree of digitalization	**+**
Driver 2: Digital Trust & Privacy Concerns	**(–)**
Driver 3: Legal Framework	**+/–**
Driver 4: Maturity of the Banking Industry	**+**
Driver 5: Transparency & Corruption	**+**
Driver 6: Economic development & Financial inclusion	**+**

Table 4: Overview of the correlation between hypothesised drivers and the level of cashlessness.

5. Discussion

Our findings support the idea that a Cashless Society is a comprehensive complex concept that involves many interactions and stakeholders as reasoned by earlier research (Arvidsson, 2019; Akinola, 2012). Systematically, the academic literature relevant for this topic acknowledges the presence of certain factors that influence the path towards being cashless. These publications intend to explain part of the phenomenon by only focusing on a particular factor or set of factors. However, none of them were aimed at providing a holistic analysis on this topic, neither from a static nor a dynamic perspective.

This thesis contributes to the understanding of the move towards a Cashless Society in three different ways. Firstly, by identifying the main drivers from previous academia research. Secondly, by statistically validating the identified drivers across a broad spectrum of countries. Lastly, by determining the direction of the validated drivers from both a static and dynamic perspective.

The review of the literature regarding the main factors that determine the level of cashlessness revealed **6 distinct drivers**: Degree of Digitalization; Digital Trust; Legal Framework; Maturity of banking industry; Transparency & Corruption, and Economic Development & Financial inclusion.

Five of these drivers were statistically validated. The only hypothesised driver that could not be statistically validated by our proposed model was "Digital Trust & Privacy Concerns". This unanticipated finding derives from the low level of significance of the Dimension 5 resulting from the regression analysis against C1. This Dimension is, as explained before, predominantly dominated by the level of digital trust based on behaviour (Variable T1). Said that, the authors of this paper decided not to discard Trust as one of the drivers due to the following reasons:

- The importance of this specific driver throughout most of the reviewed publications (Achord et al., 2017; Holst et al., 2015; Smithin, 2000).

- The fact that, despite not being statistically significant, it has the second lowest p-value and has a negative correlation against C1. The large variance of this parameter can partly explain the resulting p-value. This could mean that this driver only has impact in certain cases. For instance, this is the case of Germany, a country that it extensively known by the literature

(Achord et al., 2017) for its aversion towards becoming cashless due to its lack of trust in digital technologies despite its level of Digital and Economic Development.

- Trust is a very subjective feature hard to measure in quantitative terms, meaning that even trusted sources can deal with difficulties when evaluating this issue across countries with diverse cultural and social backgrounds.

The level of cashlessness has been analysed from two perspectives: a **static** view on the country's level of cashlessness in 2015 and a **dynamic** perspective based on the evolution of the level of cashlessness between 2010-15. Our thesis confirms that most drivers have a direct influence on both perspectives.

The first driver to be validated through our model was the: Degree of Digitalization. This driver has a positive effect on both static and dynamic perspectives. This supports the theory of Fabris (2019) and Baubeau (2016) that suggests that the progression of technology and innovation leads to an increase in digital payments and reinforces Arvidsson (2019) and Baubeau (2016) view in which new technological solutions are substituting cash in payment situations.

Legal framework was the second driver to be statistically validated. According to our model, legal framework can have either a positive or negative effect, depending on the perspective. Our literature review supports the duality of this factor (Blind, 2015). From a static perspective, the regression showed that regulatory effectiveness has a positive effect on the level of cashlessness. This finding supports Rogoff (2016), Akinola (2012) and OECD's (2002) theories that suggest that regulation implemented by governments or central banks helps to disseminate digital money faster. At the same time, C2 is negatively affected by compliance regulations. Countries that have not introduced a demanding compliance regulatory body have higher growths in terms of cashless transactions. On the other hand, it reinforces the theory of Batiz-Lazo et al. (2016) that states the rise in new forms of digital payments are linked, at least in the early stages, to a lack of regulation, weak infrastructures and poor institutional quality since it creates space for innovation and creativity to flourish. In most of the cases, such as the emergence of new forms of money or crypto currency, regulation comes after the technological novelty.

The Maturity of the Banking Industry shows a positive influence on countries' degree of cashlessness and growth. This finding is strengthened by the research of Arvidsson (2019) about Sweden, that

identifies banks as one of the necessary structures to a move towards a higher level of cashlessness. This not only supports the static perspective, but also its dynamic component. Most countries that show a progress in the development of their banking industry also display an improvement of their underlying infrastructures, which are essential to establish trustworthy and efficient payment platforms (World Economic Forum, 2015).

Transparency and Corruption are two concepts that are used interchangeably, as earlier justified. This driver, according to our regression analysis, reveals a positive effect on both parameters, C1 and C2. This fact reinforces the theoretical claims of Fabris (2019) and Warwick (2004) that argue that the elimination of physical money is considerably connected with a higher level of transparency; and also Rogoff (2016) that states that corruption prevents society from becoming cashless and delays the countries' process of phasing out cash.

The last driver to be validated through our statistical model was the Economic Development & Financial Inclusion. This driver shows a positive correlation against the static component (C1). This confirms the study of OECD (2002) that states that economies with a higher level of economic development are featured by a greater customer purchasing power with more advanced means of payments as well as a higher propensity for companies and banks to long-term investments in innovation. This supports also our Pearson Correlation Analysis that shows the high correlation (0,8) between Economic Development (E1) and the Banking Maturity (B1) variables. By contrast, this analysis reveals no statistical significance against the dynamic parameter (C2). This can be explained by the fact that the dynamic parameter has a short-term vision and, thus, contextual and cyclical economic downturns may not significantly influence the countries' trend towards becoming cashless.

As for the country clustering, which complements the main findings of our study, suggests that countries show similar patterns at different stages of cashlessness. This led to the division of countries into **4 clusters** (see figure 9). According to the cluster analysis, countries that share a high level of digital and economic progress as well as exhibiting a steady digital growth (Cluster IV) are the ones that have a greater level of cashlessness in absolute terms (C1). These countries also tend to share a good access to financial institutions, to sophisticated communications infrastructures, a higher degree of financial inclusion, greater consumer purchasing power and a higher degree of institutional stability and efficiency, as suggested by Rogoff (2016). Before reaching this final stage of cashlessness, most countries pass either through cluster II or III, depending on their progress in relative terms. The initial

and lowest stage of cashlessness (Cluster I) reveals a sharp growth in terms of digital transactions in countries with a low or medium level of digital and economic developments. This growth can be explained by the implementation of non-traditional payment methods -Saudi Arabia, Thailand and Malaysia are excellent examples- promoted by proliferation of internet and smart phones. At the same time, the cluster analysis illustrates that all the drivers are necessary to explain the cashless phenomena, including drivers such as Transparency & Corruption or Legal Framework.

To sum up, this thesis was designed to shed light on the main drivers towards a cash-free society. This thesis does not pretend, by any means, to judge this phenomenon and analyse its positive or negative effects on the overall wellbeing of societies. By having validated and measured our hypothesis and, thus, answering the initial problem statement, this thesis can be useful for different actors. In academia, other researchers can take these findings as a starting point to study deeper any of the drivers identified, to improve the statistical model development or, for instance, to analyse the consequences of a Cashless Society. Policymakers or financial institutions can also take advantage of this paper for a better understanding of this topic and, by doing that, mitigating the potential negative effects of phasing out cash.

As in any quantitative research, this paper is subjected to limitations. A first limitation is related to the generalizability of the findings. Due to the effort to represent a wide range of countries from different continents, the data collected has been constrained by its availability. Despite this limitation, the resulting research sample used for the core analysis includes countries in different stages of cashlessness. Second, the authors of this paper acknowledge that both the level of cashlessness and the hypothesized drivers can be described by alternative criteria. For instance, the level of cashlessness, as pointed out in the Methodology chapter, can be defined by the percentage of cashless transactions out of the total number of payments. For further research, these considerations should be taken into account as a way to improve the proposed model.

6. Conclusions

In the last couple of decades, the academic work around Cashless Society has significantly increased. This concept has been scrutinized from multiple angles, from anthropologic to economic perspectives. It is considered a multifaceted concept that involves complex interactions between different stakeholders and dynamics. Notwithstanding, the scientific literature regarding the subject reveals a potential knowledge gap when it comes to identifying the main drivers that lead to a Cashless Society as well as the interaction between these drivers and the resulting impact on the level of cashlessness.

This thesis ultimate purpose is to contribute vividly to the academic discussion around the drivers that explain the path towards a Cashless Society. This goal has been achieved by validating the drivers identified in the existent academic literature. In this regard, the six different hypotheses stated bellow have been developed and tested with the purpose of examining the problem statement of our thesis:

1) Degree of Digitalization
2) Digital trust & Privacy Concerns
3) Legal framework
4) Maturity of the banking industry
5) Transparency & Corruption
6) Economic development & Financial inclusion

The method followed by this paper takes a postpositivist approach as an initial point to tackle the formulated research questions. By conducting a quantitative analysis based on an original database with a large set of variables that covers countries across continents as well as the employment of different statistical tools -including PCA, clustering and regression models-, this paper has given an answer to both questions after validating each and every hypothesis.

Our study shows, backed by quantitative evidences, that there are 6 main drivers in a path towards a Cashless Society: (1) Degree of Digitalization; (2) Digital Trust; (3) Legal Framework; (4) Maturity of the Banking Industry; (5) Transparency & Corruption; and (6) Economic Development & Financial Inclusion. Three of these drivers - Degree of Digitalization; Maturity of the Banking Industry and Economic Development & Financial Inclusion- have a positive impact on the level of cashlessness;

whereas Digital Trust & Privacy has a negative effect. As for the Legal Framework, it is a double-edged sword since both positive and negative effects were registered.

This thesis contributes to the academic research in different ways. At first, it helps to clarify the existent literature by identifying the core drivers behind the path towards a Cashless Society. Then, the statistical model developed by this paper has enabled a cross-country validation of each identified driver. Finally, this same model can be used to define the direction of each driver regarding the level of cashlessness, from both a static and dynamic perspective.

Future research can build on these contributions and take this model one step further. For instance, a more elaborated model could explain the different stages of cashlessness or the consequences of each driver on the resulting level of cashlessness. At the same time, it can also encourage different stakeholders involved in the Payments Industry to acknowledge how they can be more effective in their decisions. In the meantime, it is reasonable to test our model over the coming years by merely updating the variables provided or, alternatively, by considering novel drivers based on future academic findings.

References

ABI Research & ITU-UN. (2015). *Global Cybersecurity Index & Cyberwellness Profiles.* International Telecommunications Union.

Achord, S., Chan, J., Nardani, S., & Rochemont, S. (2017). A Cashless Society: Benefits, Risks and Issues. *Institute and Faculty of Actuaries*.

Afi, D. (2013). Making sense of mobile money in urban Ghana: Personal, business, social and finanicla inclusion prospects. *IMTFI*.

Akinola, O. (2012). Cashless Society, Problems and Prospects, Data Mining Research Potentials. *International Journal of Computer Science and Telecommunications*, Vol.3. Issue 8, pp. 49-56.

Arvidsson, N. (2019). *Building a Cashless Society* (Vol. 1). Springer.

Bátiz-Lazo, B., & Efthymiou, L. (2016). Preface: News from the Cashless Front. In *The Book of Payments: Historical and Contemporary Views on the Cashless Society* (pp. 9-11). Palgrave Macmillan.

Batiz-Lazo, B., Efthymiou, L., & Michael, S. (2016). Milestones for a Global Economy. In *The Book of Payments: Historical and Contemporary Views on the Cashless Society* (pp. 391-301). Palgrave Macmillan.

Baubeau, P. (2016). Dematerialization and the Cashless society: A Look Backward, A Look Sideward. In *The Book of Payments: Historical and Contemporary Views on the Cashless Society* (pp. 85-94). Palgrave Macmillan.

Blind, B. (2015). *The Impact of Regulation on Innovation.* Manchester Institute of Innovation Research.

Broløs, A. (2016). The Future of Money. In *The Book of Payments: Historical and Contemporary Views on the Cashless Society* (pp. 367-378). Palgrave Macmillan.

Brown, B. (1997). Canada poised on brink of the cashless society. *The Halmilton Spectator*.

Camacho, J., & Ferrer, A. (2012). Cross-validation in PCA models with the element-wise k-fold (ekf) algorithm: theoretical aspects. *Journal of Chemometrics, 26*, 361-373.

Chakravorti, B., & Chaturvedi, R. S. (2017, July). *Digital Planet 2017.* The Fletcher School, Tufts University.

Chakravorti, B., Tunnard, C., & Chaturvedi, R. S. (2014). *Digital Planet: Readying for the Rise of the e-Consumer.* Institute for Business in the Global Context, The Fletcher School, Tufts University.

Chen, C., & Ganapati, S. (2018). *Is Transparency the Best Disinfectant? A Meta-Analysis on the Effect of Transparency on Government Corruption.* Steven J. Green School of International and Public Affairs, Department of Public Administration. Open Government Partnership.

Cramer, D. (2003). *Advanced Quantitative Data Analysis.* McGraw-Hill Education.

Creswell, J. W. (2014). *Research design: qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, California: SAGE Publications.

Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., & Hess, J. (2010). The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution. Washington, DC: World Bank.

Fabris, N. (2019). Cashless Society- The Future of Money or a Utopia? *Journal of Central Banking Theory and Practice*, 53-66.

FATF. (2019). *Consolidated table of assessment ratings.* FATC-GAFI. Retrieved February 21, 2019

Garcia-Swartz, D., & Layne-Farrar, A. (2006). The Move Towards a Cashless Society: Calculating the Costs and Benefits. *Review of Network Economics, 5*(2), 1-30.

Garcia-Swartz, D., Hahn, R., & Layne-Farrar, A. (2006). The Move Towards a Cashless Society: A closer Look at Payment Instruments Economics. *Review of Network Economics*, 5(2), 175-198.

Hack, U. (2016). Introduction: Mobile is the New Technology Driver. In *The Book of Payments: Historical and Contemporary Views on the Cashless Society* (pp. 283-294). Palgrave Macmillan.

Hedman, J. (2018, August 31). Going Cashless: What Can We Learn from Sweden's Experience? (Knowledge@Wharton, Interviewer, & U. o. WHARTON, Editor)

Hedman, J., & Henningsson, S. (2015). The new normal: Market cooperation in the mobile payments ecosystem. *Electronic Commerce Research and Applications, 14* (5), 305-318.

Holst, J., Hedman, J., Kjeldsen, M., & Tan, F. (2015). Payment Instrument Characteristics: A Repertory Grid Analysis. *Twenty-first Americas Conference on Information Systems.* Puerto Rico.

ISO. (2019, March). *Online Browsing Platform*. Retrieved from International Organization for Standardization: https://www.iso.org/obp/ui/#search

ITU. (2017). *Global Cybersecurity Index 2017.* International Telecommunication Union.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.

Kapron, Z., & Meertens, M. (2017). *Social Networks, E-Commerce Platforms and the Growth of Digital Payment Ecosystems in China.* Better than Case Alliance.

Kenneth, R. (2015). Costs and benefits of phasing out paper currency. *NBER Macroeconomics Annual, 29.*

Kodinariya, T. M., & Makwana, P. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal of Advance Research in Computer Science and Management Studies, 1*(6), 90-95.

Lindstedt, C., & Naurin, D. (2006). *Transparency against Corruption. A Cross-country Analysis.* Goteborg University. International Political Science Association.

Liu, Y., & Gopalakrishnan, V. (2017). An Overview and Evaluation of Recent Machine Learning Imputation Methods Using Cardiac Imaging Data. *Data*.

Maurer, B. (2016). Foreword: Friction and Fantasies of the Cashless Society. In *The Book of Payments: Historical and Contemporary Views on the Cashless Society* (pp. 5-7). Palgrave Macmillan.

OECD. (2002). *The Future of Money.* OECD Publications.

Ortiz-Ospina, E., & Roser, M. (2019). *Our Would In Data*. Retrieved from Trust: https://ourworldindata.org/trust#

Robin, A. (2015). *The customer's journey: Transforming the branch network.* ATM Marketplace: Wincor Nixdorf.

Rogoff, K. (2016). *The Curse of Cash.* Princeton University Press.

Rouse, M., & Verhoef, G. (2016). Mobile Banking in Africa: The Current State of Play. In *The Book of Payments: Historical and Contemporary Views on the Cashless Society* (pp. 233-257). Palgrave Macmillan.

Royall, R. M. (1986, November). The Effect of Sample Size on the Meaning of Significance Tests. *The American Statistician, 40*(4), 313-315.

Selamat, M. (2008). SAK 5090.

Shapiro, S. (1987). The Social Control of Impersonal Trust. *American Journal of Sociology, 93*(3).

Smith, S. (2017, October 10). The 10 most cashless countries in the world - where does the UK rank? *The Telegraph*. Retrieved April 10, 2019, from https://www.telegraph.co.uk/money/future-of-money/10-cashless-countries-world-does-uk-rank/

Smithin, J. (2000). What is money? *Routledge International Studies in Money and Banking*.

Transparency International. (2016). *Corruption Perceptions Index 2015.* Transparency International.

Warwick, D. (2004). Toward a Cashless Society. *The Futurist, 38*(4), 38-42.

World Bank. (2015). World Development Indicators.

World Bank. (2016). Global Payment Systems Survey.

World Economic Forum. (2015). *The Global Competitiveness Report 2015-2016.* World Economic Forum.

Appendices

Appendix A. Unified Dataset

A.1. Countries included in the Unified Dataset

```
##  [1] "Albania"
##  [2] "Armenia"
##  [3] "Australia"
##  [4] "Austria"
##  [5] "Bangladesh"
##  [6] "Bolivia"
##  [7] "Brazil"
##  [8] "Bulgaria"
##  [9] "Chile"
## [10] "Colombia"
## [11] "Costa Rica"
## [12] "Dominican Republic"
## [13] "Egypt"
## [14] "Estonia"
## [15] "Finland"
## [16] "France"
## [17] "Germany"
## [18] "Greece"
## [19] "Guatemala"
## [20] "Hungary"
## [21] "Ireland"
## [22] "Israel"
## [23] "Japan"
## [24] "Jordan"
## [25] "Latvia"
## [26] "Lithuania"
## [27] "Malaysia"
## [28] "Mauritius"
## [29] "Netherlands"
## [30] "New Zealand"
## [31] "Nigeria"
## [32] "Norway"
## [33] "Pakistan"
## [34] "Poland"
## [35] "Portugal"
## [36] "Republic of Korea"
## [37] "Russian Federation"
## [38] "Saudi Arabia"
## [39] "Serbia"
## [40] "Singapore"
## [41] "Slovakia"
## [42] "Slovenia"
## [43] "South Africa"
## [44] "Spain"
## [45] "Sri Lanka"
## [46] "Sweden"
## [47] "Switzerland"
## [48] "Thailand"
## [49] "Turkey"
## [50] "United Arab Emirates"
## [51] "United Kingdom of Great Britain and Northern Ireland"
## [52] "United States of America"
## [53] "Viet Nam"
## [54] "Zimbabwe"
```

A.2. Sample of values

```
##    Country or area name Code       C1         C2   D1   D2   T1    T2 R1
## 1:            Singapore  SGP 726.9535 0.05893477 3.69 2.35 3.26  0.81 92
## 2:    Republic of Korea  KOR 528.2707 0.08939742 3.68 1.42 1.73 -0.67 NA
## 3:               Norway  NOR 488.2201 0.05745779 3.79 1.73 2.80  0.39 76
## 4:          New Zealand  NZL 464.4335 0.03775956 3.54 2.38 2.66  0.15 NA
## 5:               Sweden  SWE 438.7181 0.06663803 3.79 1.09 2.52 -0.82 91
## 6:            Australia  AUS 414.4244 0.06739513 3.55 1.25 1.85 -1.05 78
##    R2       E1           E2       B1           B2 CP1          CP2
## 1: 14 5.676669  0.007162705 5.571399 -0.006475373  85 -0.017828800
## 2: NA 4.986650  0.002279708 3.602648 -0.020236310  54  0.000000000
## 3: 13 5.405860  0.010027827 5.209466 -0.005477997  88  0.004608490
## 4: NA 5.254458  0.013337444 5.728251  0.018401431  91 -0.004338558
## 5: 17 5.434574 -0.004423921 4.988134 -0.006273547  89 -0.006608509
## 6: 17 5.148697  0.001445540 5.358213 -0.003448935  79 -0.019107153
```

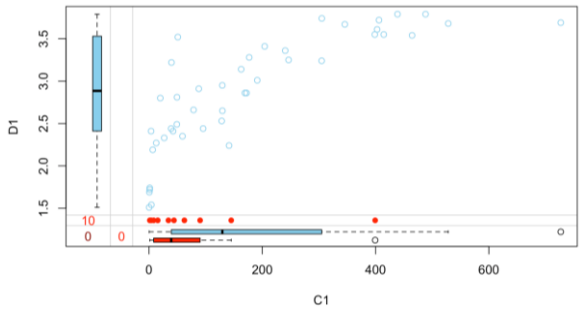
```
##    R2       E1           E2       B1           B2 CP1          CP2
## 1: 14 5.676669  0.007162705 5.571399 -0.006475373  85 -0.017828800
## 2: NA 4.986650  0.002279708 3.602648 -0.020236310  54  0.000000000
## 3: 13 5.405860  0.010027827 5.209466 -0.005477997  88  0.004608490
## 4: NA 5.254458  0.013337444 5.728251  0.018401431  91 -0.004338558
## 5: 17 5.434574 -0.004423921 4.988134 -0.006273547  89 -0.006608509
## 6: 17 5.148697  0.001445540 5.358213 -0.003448935  79 -0.019107153
##         GC1          GC2          G1           G2    CY1  CY2    CY3   CY4
## 1: 54940.86  0.033614130 3.04098e+11  0.051635279 0.6765 0.75 0.6667 0.750
## 2: 27105.08  0.041796671 1.38276e+12  0.047867845 0.7059 1.00 0.6667 0.875
## 3: 74498.14 -0.032257979 3.86663e+11 -0.020625964 0.7353 1.00 0.6667 0.750
## 4: 38649.38  0.027834324 1.77621e+11  0.039158090 0.7353 1.00 0.8333 0.875
## 5: 50812.19 -0.004902503 4.97918e+11  0.003876808 0.6471 0.75 0.6667 0.625
## 6: 56561.41  0.017205883 1.34903e+12  0.033474097 0.7647 0.75 0.6667 0.875
##      CY5   CY6   CY7   CY8     DEI1         DEI2        F1           F2
## 1: 0.750 0.500 0.925 0.925 56.20681  0.020544836 0.9635260 -0.006384444
## 2: 0.625 0.500 0.782 0.782 50.62514  0.020865828 0.9436082  0.004685136
## 3: 0.875 0.500 0.786 0.786 47.03921 -0.012803597 1.0000000           NA
## 4: 0.625 0.500 0.718 0.718 46.25875  0.011793060 0.9952554  0.000297605
## 5: 0.625 0.625 0.733 0.733 55.23459 -0.007945216 0.9971992  0.002445290
## 6: 0.875 0.625 0.824 0.824 48.14280 -0.005647665 0.9885957 -0.000691198
##           F3
## 1: 0.8702403
## 2: 0.8610544
## 3: 0.9836910
## 4: 0.9761475
## 5: 0.9921191
## 6: 0.9490806
```

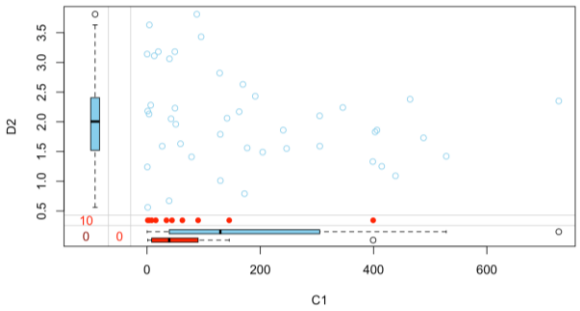
A.3. Variables Description

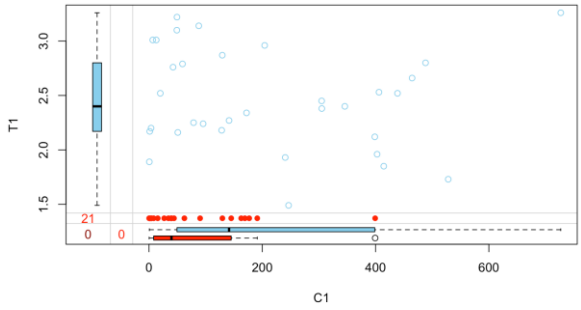
```
##  Country or area name     Code                C1
##  Length:54            Length:54          Min.   :  0.236
##  Class :character     Class :character   1st Qu.: 28.663
##  Mode  :character     Mode  :character   Median : 92.865
##                                          Mean   :162.751
##                                          3rd Qu.:244.940
##                                          Max.   :726.953
##
##        C2                 D1              D2              T1
##  Min.   :-0.08062   Min.   :1.510   Min.   :0.560   Min.   :1.490
##  1st Qu.: 0.05783   1st Qu.:2.410   1st Qu.:1.535   1st Qu.:2.170
##  Median : 0.10460   Median :2.885   Median :2.005   Median :2.400
##  Mean   : 0.20620   Mean   :2.876   Mean   :2.041   Mean   :2.459
##  3rd Qu.: 0.22731   3rd Qu.:3.525   3rd Qu.:2.393   3rd Qu.:2.800
##  Max.   : 1.62595   Max.   :3.790   Max.   :3.810   Max.   :3.260
##                     NA's   :10      NA's   :10      NA's   :21
##        T2              R1              R2              E1
##  Min.   :-1.05   Min.   : 44.0   Min.   : 2.00   Min.   :3.445
##  1st Qu.:-0.49   1st Qu.: 75.0   1st Qu.:10.25   1st Qu.:4.216
##  Median :-0.04   Median : 80.0   Median :12.00   Median :4.472
##  Mean   : 0.01   Mean   : 81.6   Mean   :13.13   Mean   :4.600
##  3rd Qu.: 0.42   3rd Qu.: 91.0   3rd Qu.:16.75   3rd Qu.:5.143
##  Max.   : 1.10   Max.   :106.0   Max.   :23.00   Max.   :5.759
##  NA's   :21      NA's   :24      NA's   :24
```

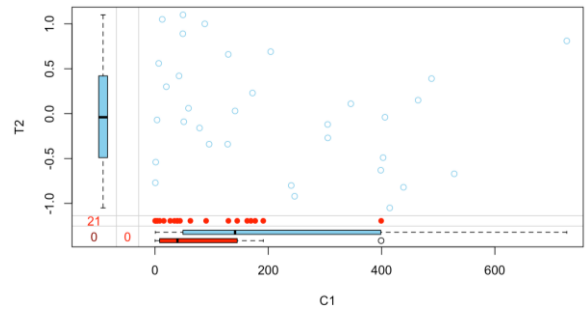
```
##        E2                   B1              B2                 CP1
##  Min.   :-0.0174245   Min.   :2.813   Min.   :-0.066487   Min.   :21.00
##  1st Qu.: 0.0009448   1st Qu.:3.614   1st Qu.:-0.021989   1st Qu.:37.25
##  Median : 0.0043331   Median :4.236   Median :-0.010794   Median :53.50
##  Mean   : 0.0040359   Mean   :4.212   Mean   :-0.011043   Mean   :55.61
##  3rd Qu.: 0.0067501   3rd Qu.:4.689   3rd Qu.: 0.001717   3rd Qu.:73.75
##  Max.   : 0.0268173   Max.   :5.728   Max.   : 0.031418   Max.   :91.00
##
##       CP2                GC1             GC2
##  Min.   :-0.02635   Min.   : 1210   Min.   :-0.076605
##  1st Qu.:-0.00446   1st Qu.: 5769   1st Qu.:-0.008856
##  Median : 0.01200   Median :13688   Median : 0.018163
##  Mean   : 0.01234   Mean   :22243   Mean   : 0.015712
##  3rd Qu.: 0.02624   3rd Qu.:38140   3rd Qu.: 0.035756
##  Max.   : 0.06668   Max.   :82016   Max.   : 0.098177
##
##        G1                  G2                CY1              CY2
##  Min.   :1.055e+10   Min.   :-0.08165   Min.   :0.0882   Min.   :0.0000
##  1st Qu.:6.503e+10   1st Qu.:-0.01126   1st Qu.:0.2941   1st Qu.:0.5000
##  Median :2.920e+11   Median : 0.02811   Median :0.5147   Median :0.7500
##  Mean   :9.061e+11   Mean   : 0.02455   Mean   :0.4798   Mean   :0.6991
##  3rd Qu.:6.152e+11   3rd Qu.: 0.05042   3rd Qu.:0.6765   3rd Qu.:1.0000
##  Max.   :1.812e+13   Max.   : 0.11094   Max.   :0.8235   Max.   :1.0000
##
##       CY3              CY4              CY5              CY6
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.3333   1st Qu.:0.2500   1st Qu.:0.2500   1st Qu.:0.1250
##  Median :0.5000   Median :0.5625   Median :0.5000   Median :0.5000
##  Mean   :0.4661   Mean   :0.5069   Mean   :0.4653   Mean   :0.3681
##  3rd Qu.:0.6667   3rd Qu.:0.7500   3rd Qu.:0.6250   3rd Qu.:0.5000
##  Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :0.6250
##
##       CY7              CY8              DEI1            DEI2
##  Min.   :0.1140   Min.   :0.1140   Min.   :13.77   Min.   :-0.025845
##  1st Qu.:0.4260   1st Qu.:0.4260   1st Qu.:28.78   1st Qu.:-0.007870
##  Median :0.5800   Median :0.5800   Median :39.39   Median : 0.005944
##  Mean   :0.5724   Mean   :0.5724   Mean   :37.93   Mean   : 0.009005
##  3rd Qu.:0.7552   3rd Qu.:0.7552   3rd Qu.:47.15   3rd Qu.: 0.020858
##  Max.   :0.9250   Max.   :0.9250   Max.   :56.21   Max.   : 0.066466
##                                    NA's   :16      NA's   :16
##        F1               F2                  F3
##  Min.   :0.1304   Min.   :-0.0652084   Min.   :0.07412
##  1st Qu.:0.5825   1st Qu.: 0.0006656   1st Qu.:0.34113
##  Median :0.8144   Median : 0.0167737   Median :0.61479
##  Mean   :0.7316   Mean   : 0.0451382   Mean   :0.60866
##  3rd Qu.:0.9663   3rd Qu.: 0.0883215   3rd Qu.:0.90247
##  Max.   :1.0000   Max.   : 0.2281883   Max.   :0.99212
##                   NA's   :2
```

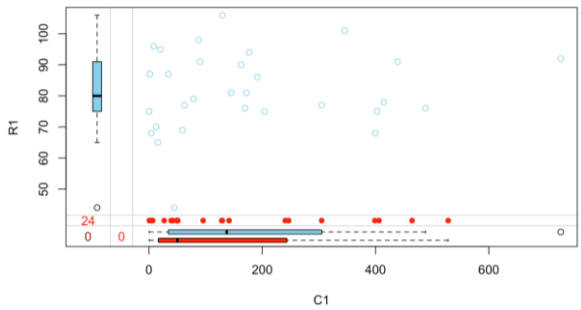
A.4. Distribution of missing values by variable against C1

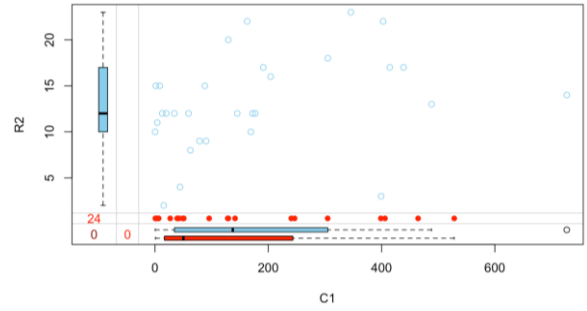


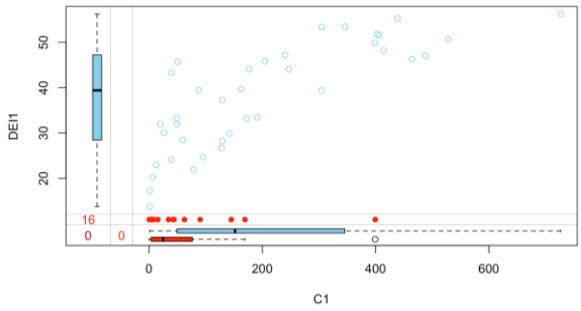


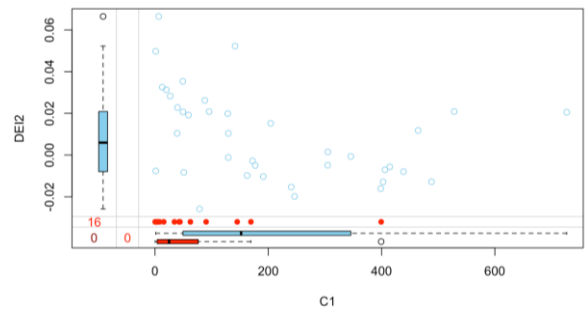


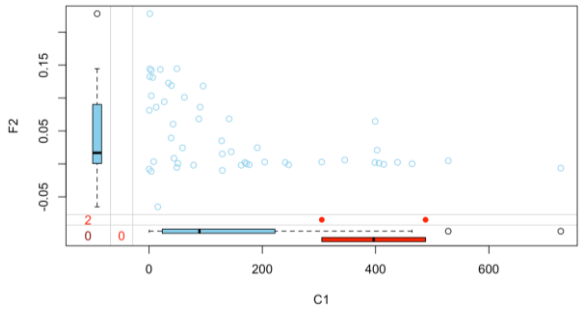












Appendix B. Research samples

B.1. Countries included in the Research sample with Complete Data

```
##  [1] "Australia"
##  [2] "Colombia"
##  [3] "Hungary"
##  [4] "Ireland"
##  [5] "Malaysia"
##  [6] "Saudi Arabia"
##  [7] "Singapore"
##  [8] "Slovenia"
##  [9] "Spain"
## [10] "Sweden"
## [11] "Thailand"
## [12] "United Kingdom of Great Britain and Northern Ireland"
## [13] "United States of America"
```

B.2. Countries included in the Optimised Research sample

```
##  [1] "Australia"
##  [2] "Austria"
##  [3] "Bangladesh"
##  [4] "Bolivia"
##  [5] "Brazil"
##  [6] "Bulgaria"
##  [7] "Chile"
##  [8] "Colombia"
##  [9] "Egypt"
## [10] "Estonia"
## [11] "Finland"
## [12] "France"
## [13] "Germany"
## [14] "Greece"
## [15] "Hungary"
## [16] "Ireland"
## [17] "Israel"
## [18] "Japan"
## [19] "Jordan"
## [20] "Latvia"
## [21] "Malaysia"
## [22] "Netherlands"
## [23] "New Zealand"
## [24] "Nigeria"
## [25] "Norway"
## [26] "Pakistan"
## [27] "Poland"
## [28] "Portugal"
## [29] "Republic of Korea"
## [30] "Russian Federation"
## [31] "Saudi Arabia"
## [32] "Singapore"
## [33] "Slovakia"
## [34] "Slovenia"
## [35] "South Africa"
## [36] "Spain"
## [37] "Sweden"
## [38] "Switzerland"
## [39] "Thailand"
## [40] "Turkey"
## [41] "United Arab Emirates"
## [42] "United Kingdom of Great Britain and Northern Ireland"
## [43] "United States of America"
## [44] "Viet Nam"
```

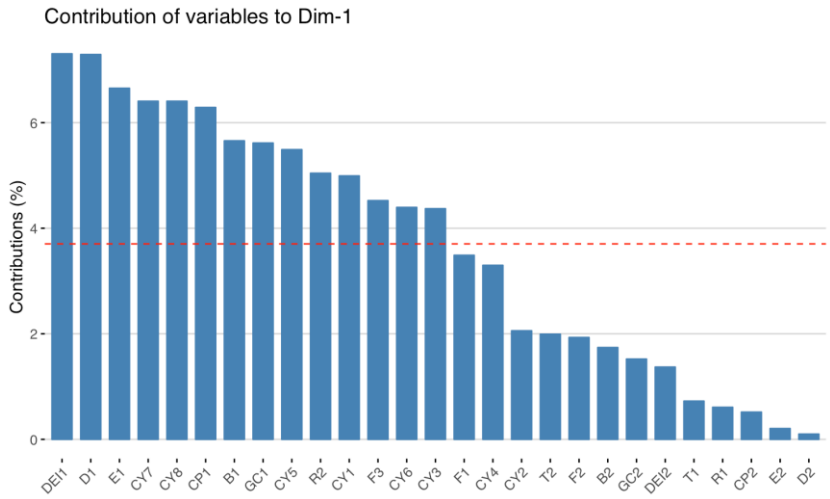
Appendix C. Data analysis methods

C.1. Overview of the PC analysis (RS with Complete Data)

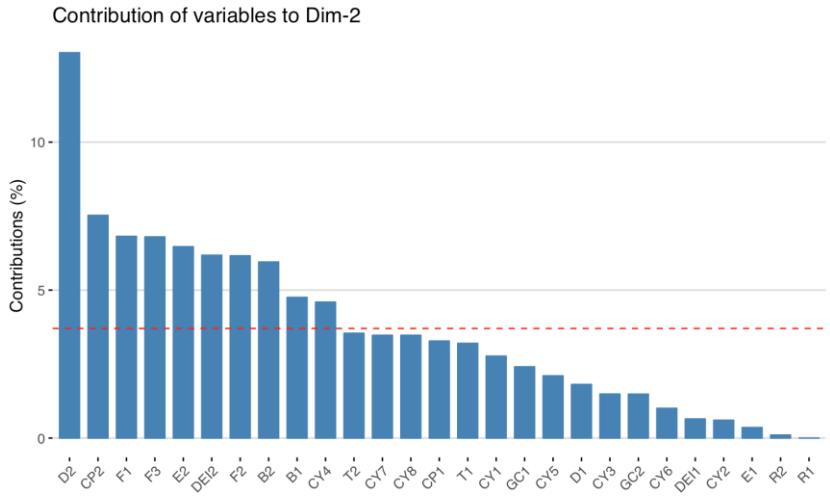
```
##          eigenvalue percentage of variance cumulative percentage of variance
## comp 1  10.59346743             39.2350646                           39.23506
## comp 2   6.09815847             22.5857721                           61.82084
## comp 3   3.93189891             14.5625886                           76.38343
## comp 4   2.32047296              8.5943443                           84.97777
## comp 5   1.28135452              4.7457575                           89.72353
## comp 6   1.10699747              4.0999906                           93.82352
## comp 7   0.73983309              2.7401225                           96.56364
## comp 8   0.36286239              1.3439348                           97.90757
## comp 9   0.25141595              0.9311702                           98.83875
## comp 10  0.14823436              0.5490161                           99.38776
## comp 11  0.08717699              0.3228777                           99.71064
## comp 12  0.07812746              0.2893610                          100.00000
```

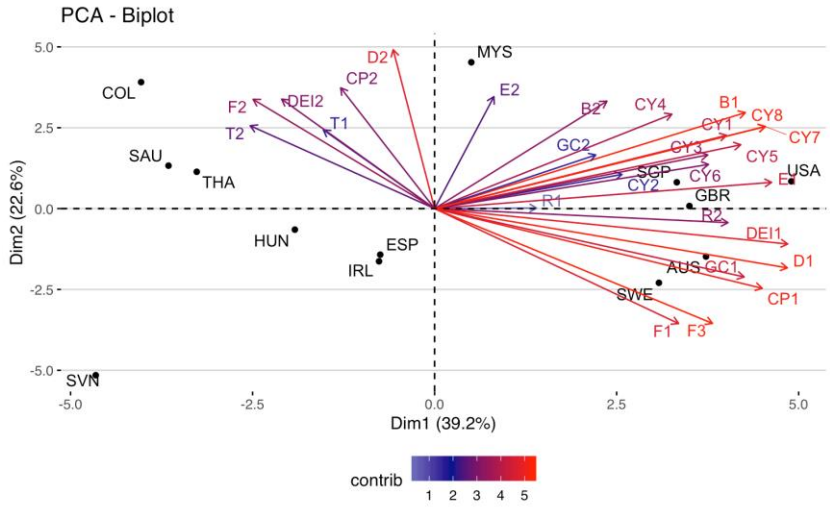
C.2. Contribution graphs for the different components (RS with Complete Data)

Dimension 1:

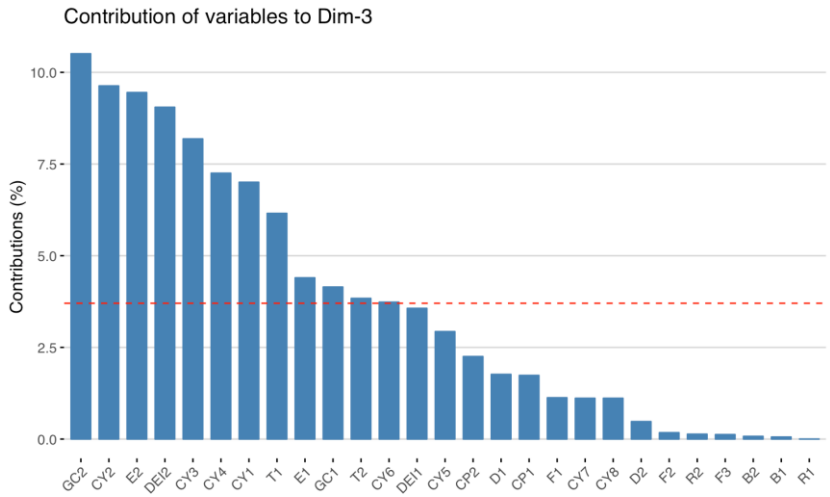


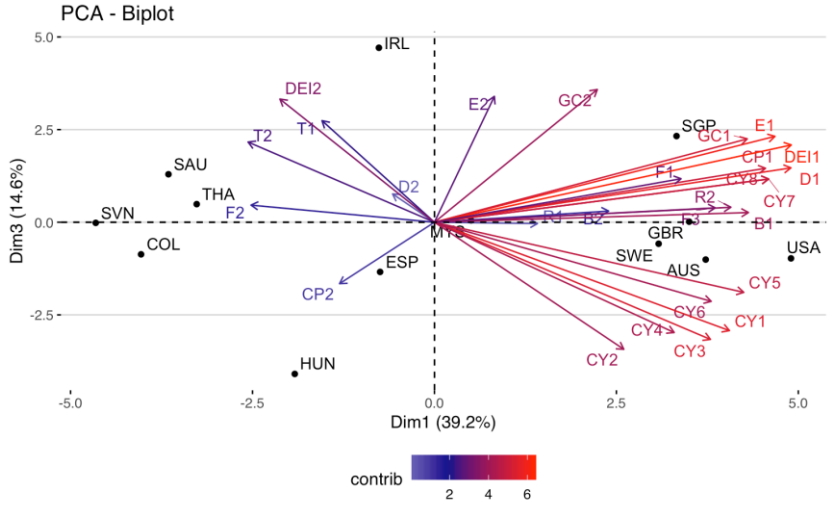
Dimension 2:



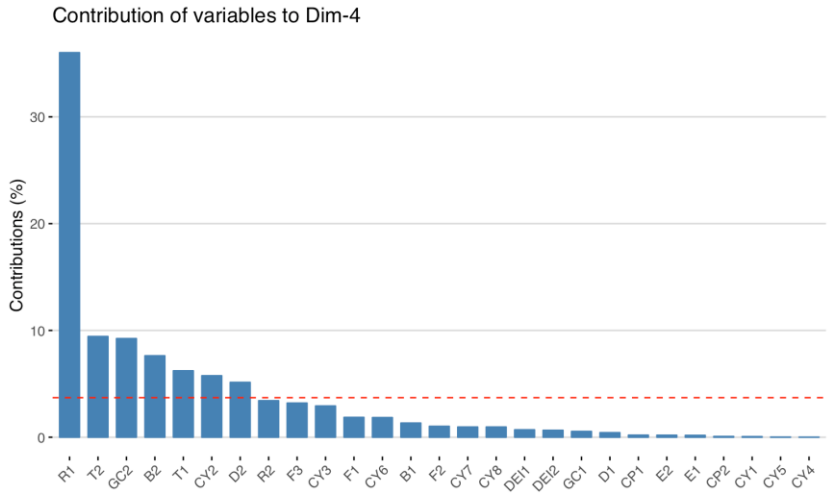


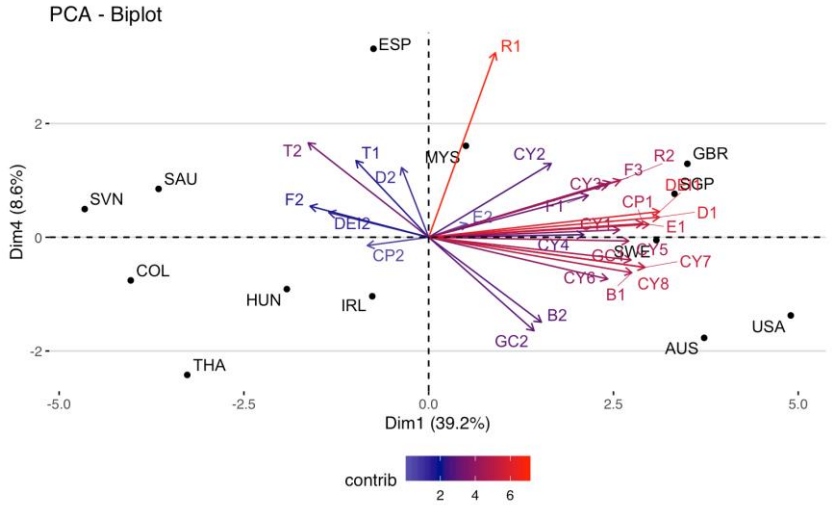
Dimension 3:



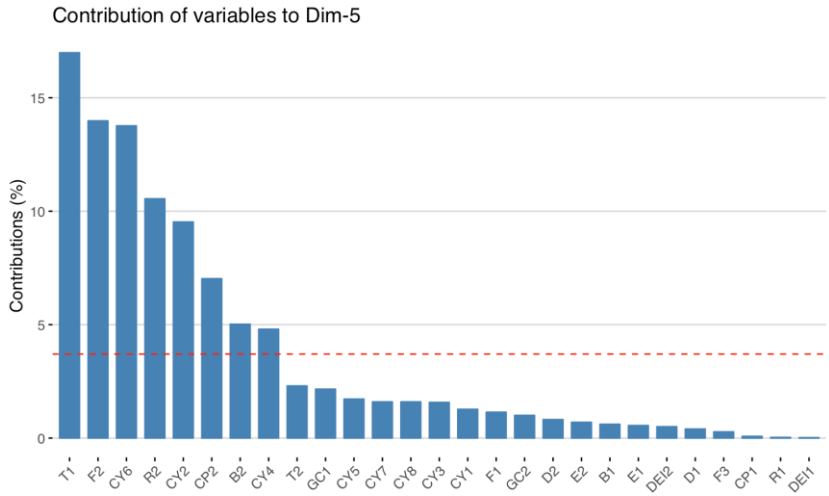

PCA - Biplot
IRL
DEI2
E2
GC2
SGP
E1
T2
T1
GC1
CP1
DEI1
SAU
D2
F1
CY8
D1
THA
R2
CY7
SVN
F2
R1
B2
F3
B1
MYS
GBR
COL
USA
ESP
SWE
AUS
CP2
CY5
CY6
CY1
CY4
CY3
CY2
HUN
Dim3 (14.6%)
Dim1 (39.2%)
contrib

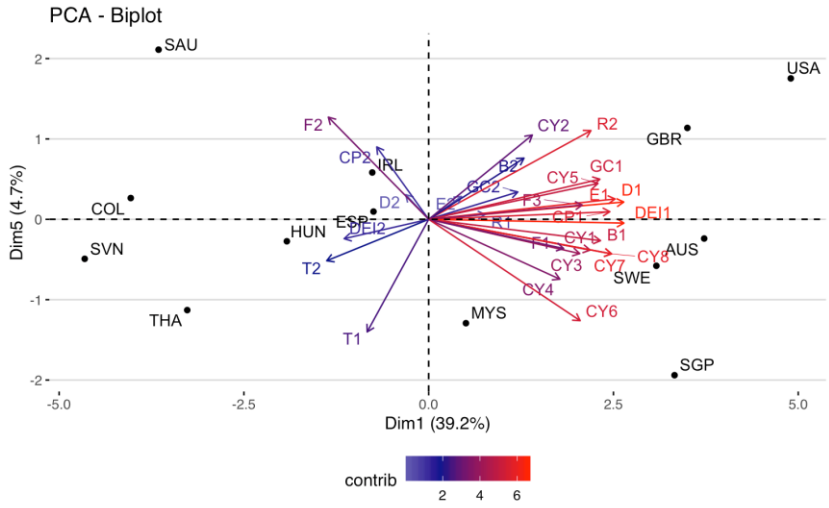
Dimension 4:


Contribution of variables to Dim-4
Contributions (%)
R1
T2
GC2
B2
T1
CY2
D2
R2
F3
CY3
F1
CY6
B1
F2
CY7
CY8
DEI1
DEI2
GC1
D1
CP1
E2
E1
CP2
CY1
CY5
CY4

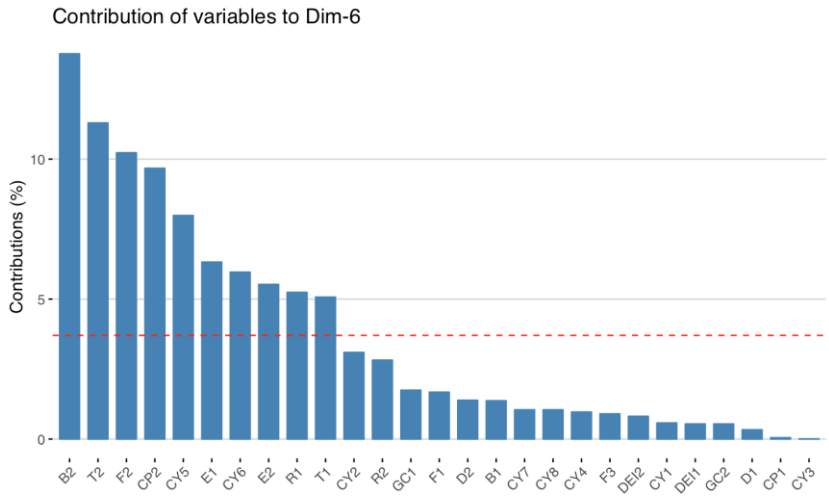

PCA - Biplot
ESP
R1
T2
T1
MYS
CY2
R2
GBR
F3
SAU
CY3
SGP
SVN
D2
CP1
F2
F1
D1
DEI2
E2
CY1
E1
CY4
SWE
CY5
CP2
GC1
CY7
COL
CY6
CY8
B1
HUN
IRL
B2
USA
GC2
AUS
THA
Dim4 (8.6%)
Dim1 (39.2%)
contrib

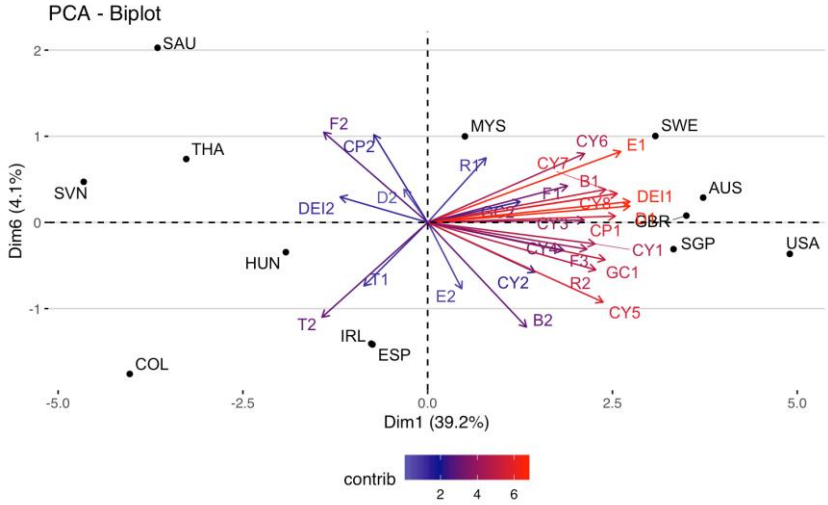
Dimension 5:



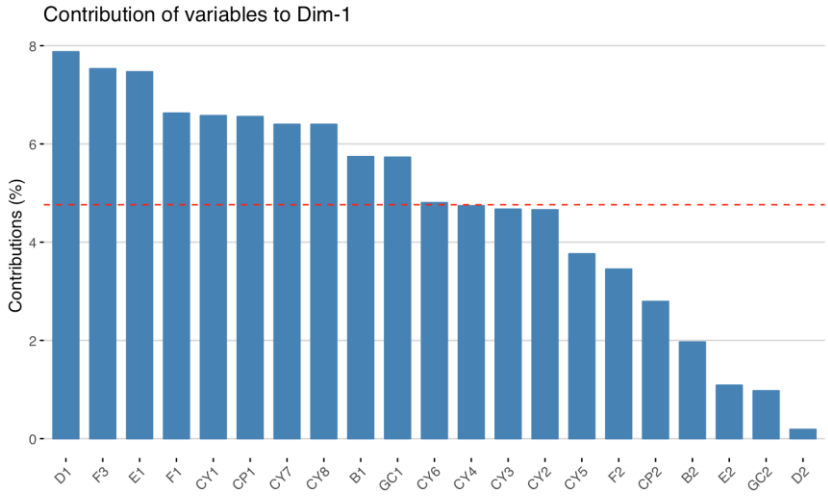


Dimension 6:

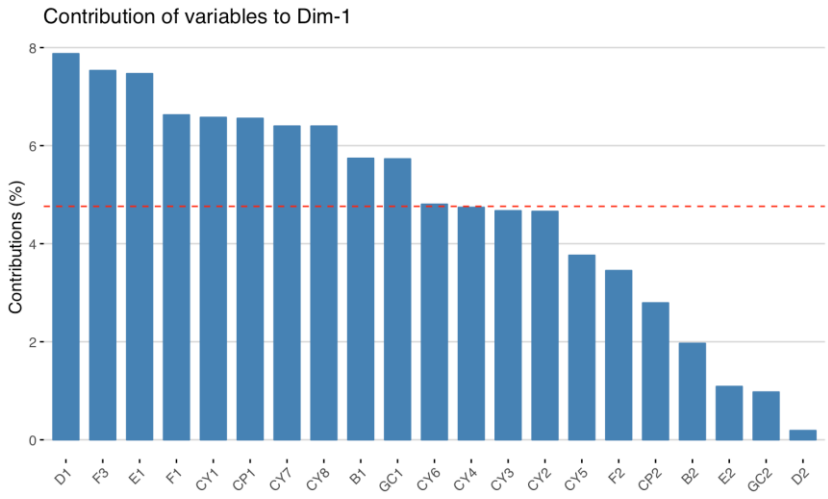


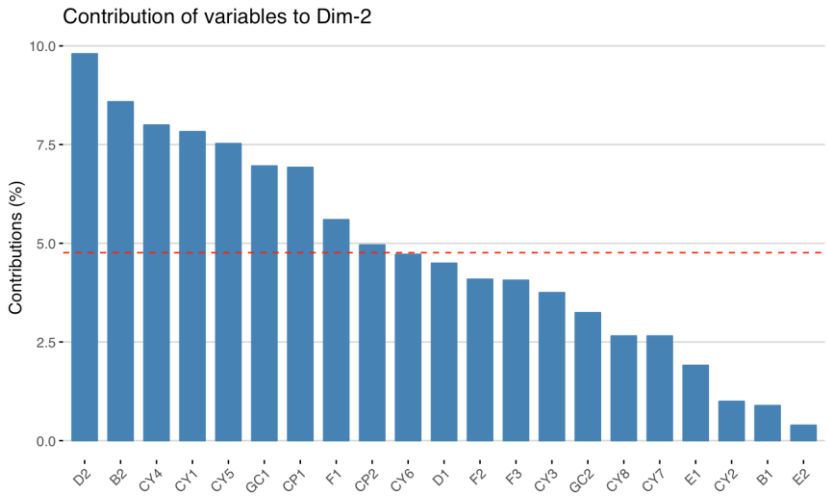


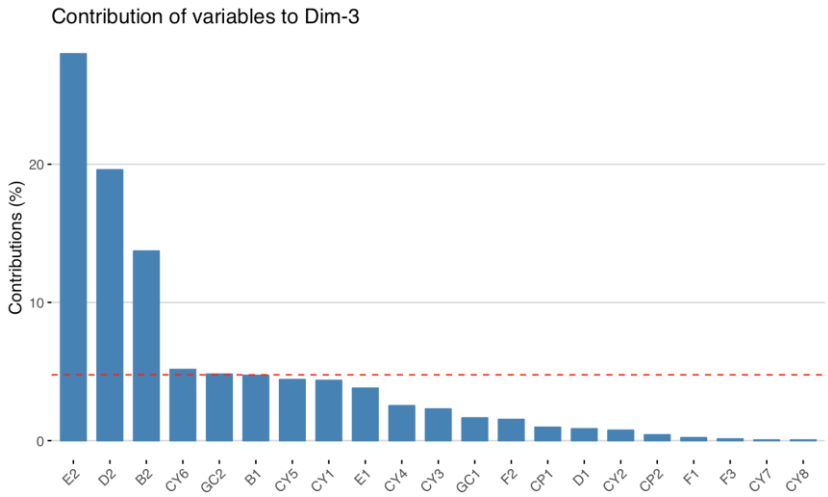
C.3. Overview of the PC analysis (Optimised RS)



Contribution graphs for the main components (Optimised RS)

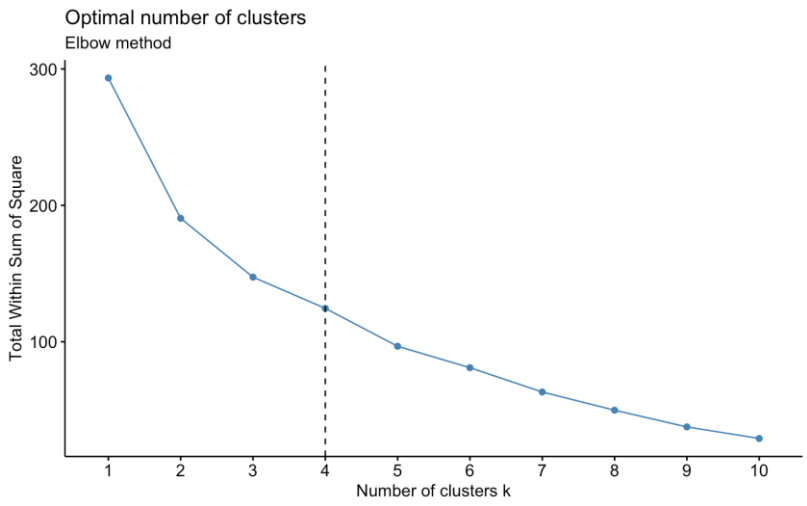




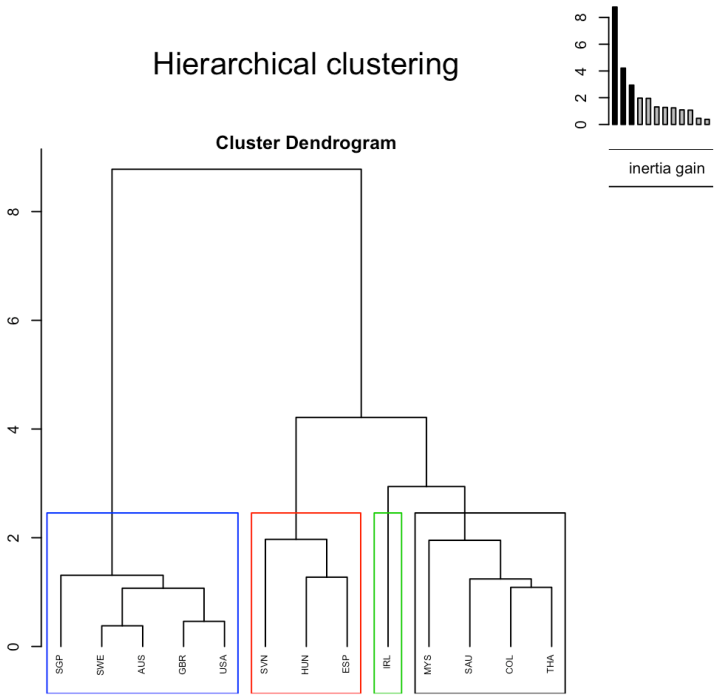


C.4. Cluster analysis (RS with Complete Data)

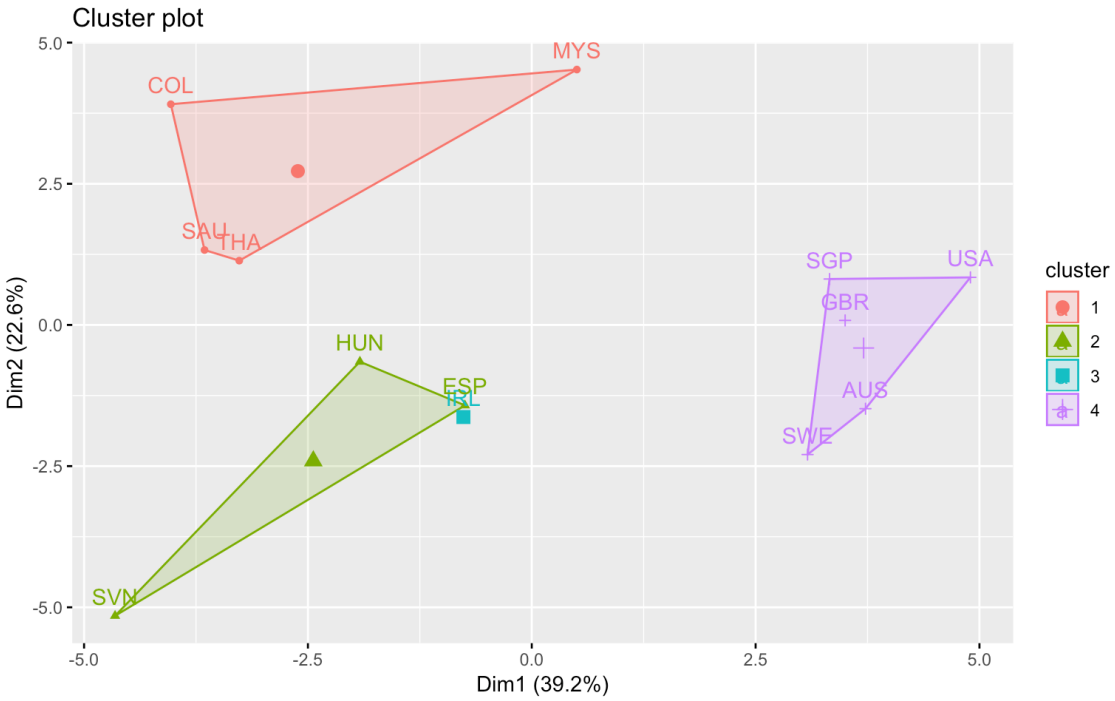
Optimal number of clusters using the Elbow Method



Cluster Dendogram

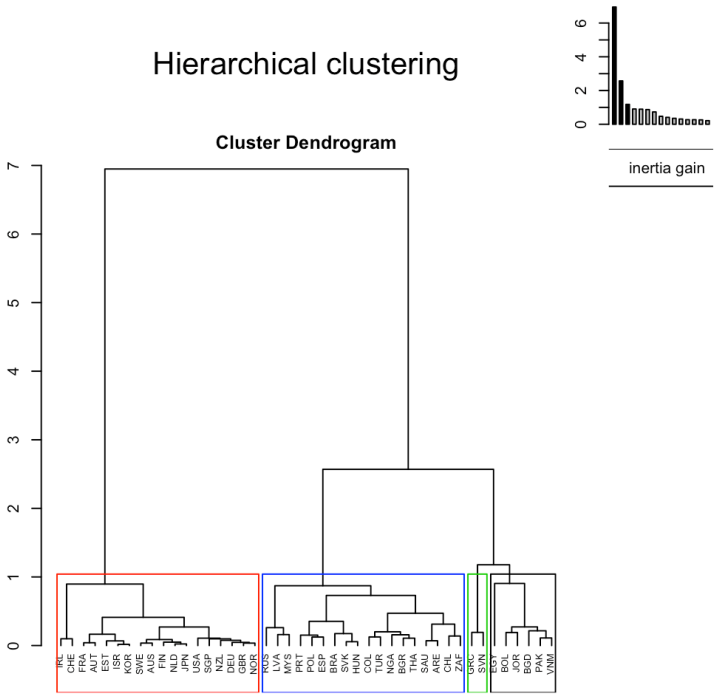


Cluster plot against Dim-1 and Dim-2

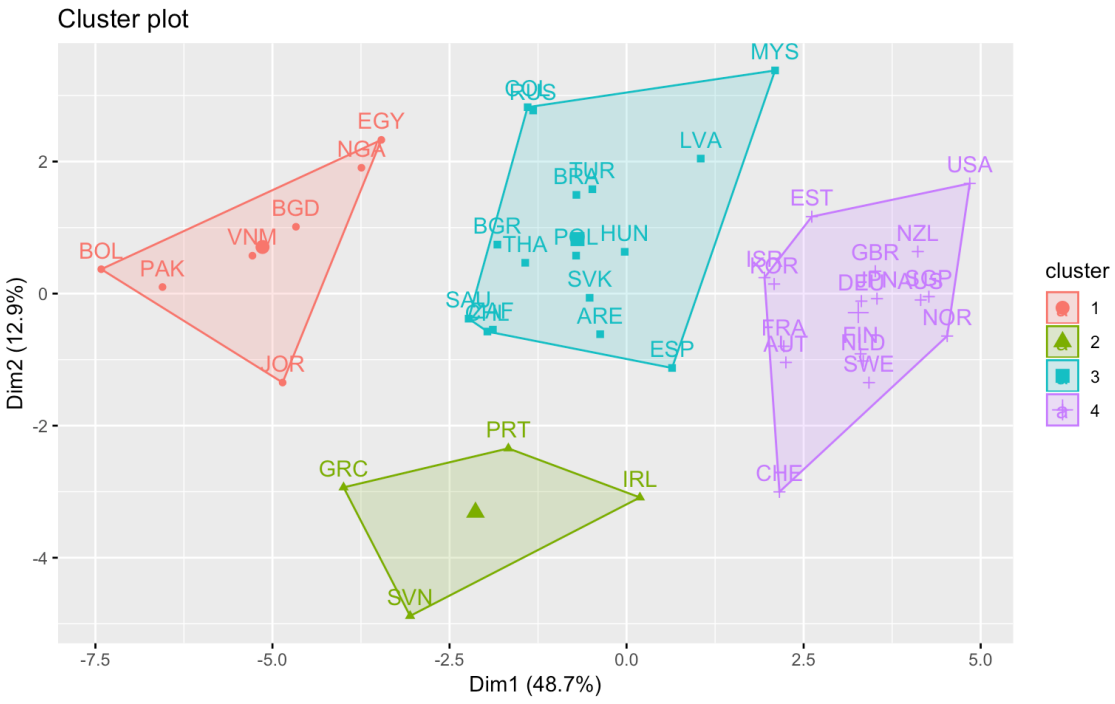


C.5. Cluster analysis (Optimised RS)

Cluster Dendogram



Cluster plot against Dim-1 and Dim-2



C.6. Regression analysis (RS with Complete Data)

C1 as response variable:

```
##
## Call:
## glm(formula = C1 ~ ., data = as.data.frame(dimensions))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -139.35   -56.03   -12.11    49.50   194.37
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  237.878     34.682   6.859 0.000473 ***
## Dim.1         50.970     10.656   4.783 0.003053 **
## Dim.2        -19.889     14.044  -1.416 0.206501
## Dim.3         22.338     17.491   1.277 0.248748
## Dim.4         -2.371     22.768  -0.104 0.920441
## Dim.5        -46.430     30.639  -1.515 0.180448
## Dim.6         -3.697     32.963  -0.112 0.914356
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 15636.95)
##
##     Null deviance: 544730  on 12  degrees of freedom
## Residual deviance:  93822  on  6  degrees of freedom
## AIC: 168.39
##
## Number of Fisher Scoring iterations: 2
```

C2 as response variable

```
##
## Call:
## glm(formula = C2 ~ ., data = as.data.frame(dimensions))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.0868  -1.1373   0.3001   0.8327   2.9616
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   8.3607     0.6984  11.971 2.06e-05 ***
## Dim.1        -0.3908     0.2146  -1.821  0.11844
## Dim.2         1.2182     0.2828   4.307  0.00505 **
## Dim.3         0.3885     0.3522   1.103  0.31226
## Dim.4        -1.1973     0.4585  -2.611  0.04004 *
## Dim.5         0.8632     0.6170   1.399  0.21131
## Dim.6         2.6554     0.6638   4.000  0.00712 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 6.341129)
##
##     Null deviance: 341.567  on 12  degrees of freedom
## Residual deviance:  38.047  on  6  degrees of freedom
## AIC: 66.853
##
## Number of Fisher Scoring iterations: 2
```