In Search For A Place In World Order:

Dynamic Modelling of National Power Using A Systems Approach

R. (Rémon) ten Bhömer

Technische Universiteit Delft

In Search For A Place In World Order:

Dynamic Modelling of National Power Using A Systems Approach

by

R. (Rémon) ten Bhömer

to obtain the degree of Master of Science
at the Delft University of Technology,
Msc. Programme Engineering, and Policy Analysis

Student number:	4388658	
Project duration:	February 2018 – August 2018	
Thesis committee:	dr. E. (Erik) Pruyt,	TU Delft, supervisor
	Prof. dr. M. J. (Jeroen) van den Hoven,	TU Delft, chair

TU Delft

Preface

I would like to thank my supervisors dr. Erik Pruyt and Prof.dr. Jeroen van den Hoven for all their support during this thesis. I would also like to thank Scott Cunningham in providing me with helpful insights on statistical analysis. Finally, I would like to thank my fellow students for listening to me and helping me structuring my thoughts in the process of writing this thesis.

R. (Rémon) ten Bhömer
Delft, August 15

Abstract

Difficulties arise in the making of predictions, since ideas and views on the future are often based on current conditions and, with only minor adjustments, it is simply expected that those conditions will continue into the future. With the goal of supporting experts in their scenario analysis this thesis is part of a pilot between the analysis network on national security (Analistennetwerk Nationale Veiligheid) of the Netherlands and the Technical University Delft. It explores if system dynamics is an useful approach in the modelling and simulation of changes in state power dynamics. It does so by researching literature on national power and through the analysis multi-scale geopolitical model to simulate over 217 countries. Exploratory modelling and analysis has been performed on the power dynamics of The United States, China, and Russia. This thesis concludes that system dynamics is an appropriate approach in the modelling and simulation of geopolitics and shifts in power dynamics.

But, that uncertainties in the relationships between variables, missing data, and the modelling choices made to overcome these issues make it very difficult to build a good causal model to support expert scenario analysis.

Contents

List of Figures

List of Tables

Biases of the modeller

"It is better to state your biases than to pretend you don't have any" (Meadows, Richardson, & Bruckmann, 1982). In their book they state their biases. I will do this as well, as I believe this is an underestimated factor of modelling of the author:

- Is white, male, Dutch, and Msc. level academic;
- Has a bachelors in Construction Engineering, but has always been drawn to communication and optimisation;
- Is – while writing this thesis – drawn into clinical psychology and in search for his political ideas;
- Is a firm believer of the use of behaviouralism;
- Believes that humans are egocentric.

1
Introduction

Our world faces important grand challenges, such as: climate change, food scarcity, and non-renewable resource scarcity. As a society we ought to limit the development and impact of these challenges, but this is easier said then done. Uncertainty about the future, not only for our planet, but also the position of nations in world order puts pressure on the willingness to face these challenges.

Moreover, most grand challenges are also 'wicked problems': social policy problems that cannot be captured by one definition (Rittel & Webber, 1973). 'Super wicked' problems have four extra key features compared to wicked problems: time is running out, those who cause the problem also seek to provide a solution, the central authority that needs to address those problems is weak or non-existent, and irrational discounting occurs that pushes responses into the future (Levin, Cashore, Bernstein, & Auld, 2012). All these features are present in the earlier named grand challenges.

We have to face the challenges together. That is also true for keeping the world a stable place. Federica Mogherini, High Representative of the European Union for Foreign Affairs and Security Policy wants to give us (Mogherini, 2017). Stating:

"We all have the impression to live in times of global disorder. Events in the Middle East or on the Korean Peninsula can destabilise the whole world easily. Therefore, we have a common responsibility to keep a stable world."

Similarly Levin et al. (2012) state the importance of coalition building on the one hand, and norms and values on the other in their article on facing climate change. Long-term policies should be the key to solving international problems. During the fifth meeting of the Council of Councils the main mindset was long-term cooperation on these challenges in a time of changing world order and geopolitical rivalry (Mullan, 2016)[1].

However, international cooperation is easier said than done in times of rivalry. Failure to deliver and the feeling of a lack of political will can have a huge impact on future cooperation with other parties. For example, the failure of the European Union (EU) to deliver on the KEDO (Korean Peninsula Energy Organisation) cost them a spot at the table during the Six Party Talks with North Korea in 2001 and led to an Asian view on Europe being an actor of little importance or vigour (Lee, 2017). Levin et al. (2012) state as well that decision makers value our long-term goals higher than immediate interests, making addressing these challenges even more difficult.

Although more certainty about the future is desirable, it is hard to make predictions, since ideas and views on the future are often based on current conditions and, with only minor adjustments, it is simply expected that those conditions will continue into the future. People tend to ignore the concept of time, visions of the near future are much more detailed than the distant future (Young C. Kim, 2013). Likewise, Jeremey Bentham,

[1]The Council of Councils is an initiative of the Council of Foreign Relations connecting foreign policy institutes from around the world in a common conversation on issues of global governance and multilateral cooperation. It consists of twenty-five countries, roughly the G20 (BC Patient Safety and Quality Council, 2018).

the head of Shell Scenarios, points out that we "tend to be trapped within the limits of our own experience, which leads to all kinds of problems". We can only really imagine the future as being similar to what we already know from the past (Shell Global, 2018).

An example of our limited ability of qualitative forecasting could be reviewed from the literature on scenario development for Korean unification. This mainly focusses on four scenarios: (1) North Korea loses a war; (2) North-Korean people revolve the government; (3) Negotiated unification; (4) Living peacefully next to each other (Pollack, 2001; Young C. Kim, 2013). Impacts of these scenarios are mainly described on state level or regional level, and based on similar historical situations like the German unification or the Vietnamese unification (Brookings Institution, 2017; Huh et al., 2012; Silbereisen, 2016; Vogel & Best, 2016). Some research focusses only on one of the scenarios and its impact on global economy (Martin, 2017a, 2017b). Although this research is highly relevant to get an understanding of the problem and the system, there is a large part of possible dynamics missing as the system becomes too complex to grasp qualitatively.

The European Union (EU) and the Netherlands have been searching for over a decade on how to position themselves in a changing world order. The EU has done this through a project called "GR:EEN" researching "Europe's role in the emerging global order" (GR:EEN, 2014). The research advice board for government policy of the Netherlands (WRR: Wetenschappelijke Raad voor Regeringsbeleid) hosted a conference in 2011 on *power shifts in a changing world order: the role of the EU and the position of the Netherlands* (European Parliament, 1997). They state that the western world has two main challenges. Firstly, the emergence of Asia in the international arena. Secondly, that there will not be a shift in power, but a diffusion in power instead. They ask themselves the question: "Will the EU become a victim of a global 'struggle' for resources? Or will it lead the way in terms of a transition to a more sustainable world?" (European Parliament, 1997).

Recently, the analysis network on national security (Analistennetwerk Nationale Veiligheid, ANV) of The Netherlands asked the Delft University of Technology (TU Delft) to explore the possibilities of modelling autonomous developments using a systems approach. ANV identifies five categories of autonomy developments: (1) ecological (climate-change and environmental pressure), (2) demographic-social (demographic population composition, but also societal values and norms), (3) international-political (shifting power dynamics, tension increase between powers, and regional instability), (4) international-economic (economic inequality), and (5) technological (technological change on society) (Analistennetwerk Nationale Veiligheid, 2016). In a specific case on geopolitics the ANV asked the TU Delft to explore and simulate, based on components of power, the changes in power dynamics, the impact on key regions in the world, and the resulting impacts on The Netherlands.

This thesis is a first step in satisfying that demand by investigating whether it is possible to model geopolitics so that it identifies national power shifts.Therefore, this thesis positions itself on the third autonomous development: international-political. Five questions will be addressed: (1) What is national Power? (2) What components does national power consist of? (3) Are these components possible to model? (4) How can these components be modelled? And finally, (5) is system dynamics a useful modelling approach in the modelling of geopolitics? Background knowledge on national power and important concepts was gathered by reviewing literature and expert consolidation.

Due to the exploratory objective of this thesis and the fact that national power overlaps several research fields, this thesis is only a first step. It tries to identify what should be included in a geopolitical simulation model after which such a model is built.

This research fits seamlessly in the 'Engineering, and Policy Analysis' Masters programme at the Technical University Delft in which complex, high actor involvement challenges are aimed for to be structured. By researching shifts in national power by the use of modelling and simulation, this challenge is addressed on the highest level, providing exploratory scenario analysis as a basis for durable policy development.

This thesis is structured as followed. A literature review of geopolitics and national power is provided in chapter 2. Whereafter, in chapter 3 methodologies are described. The modelling of power, together with the identified and built model structures is discussed in Chapter 4. In chapter 5 the model results are discussed, as well as the model itself. Conclusions are provided in chapter 6 together with a discussion and suggested further research.

2

On Geopolitics and National Power

The following chapter provides a literature review on two main concepts for the measurement of national power. First, we discuss the changing definition of geopolitics, define national power for the purpose of this thesis, and review national power in the context of its components and measurement.

2.1. Geopolitics, a transforming definition

Over the years many definitions of geopolitics have been coined and used in various fields of study. As definitions partly set the scope of a research this section will first shortly summarise the history of the definition, followed by an explanation of the definition that will be used in the rest of this thesis of geopolitics.

For the ANV, geopolitics is the influence of geographical factors on international political issues. More specifically, the struggle in constraining land-, sea-, and airspace borders of political influence (Analistennetwerk Nationale Veiligheid, 2016). However, this definition seems insufficient for the issues it tries to address.

In 1928 Haushofer wrote "Geopolitics is the science of conditioning of political processes by the earth." (Herwig, 1999). Sahoo (2010) provides us with a clearer definition: "Geopolitics views a states political position in the world on the basis of geographical context. It analyses the space, location, size, and resources of nation states."

Rudolf Kjellen first used the term *geopolitik* (geopolitics) (Kjéllen, 1916). Spykman writes that the geographic being of the state is the base from which it operates in peace and war. Geography is the most fundamental conditioning factor in the formulation of national policy, because it is most permanent. For Morgenthau (1904-1980) geography is one of the dominant factors which decide national power (Morgenthau, 2001). For him geopolitics (geopolitik) is a struggle for power, so power is vital for national survival.

In Germany geography was forced-replaced by geopolitik, and during the 1930s it got a direct association with the intellectual apparatus of the Third Reich. Due to these circumstances geopolitik was 'banned' from the intellectual world, but it survived as applied political geography (Kasperson & Minghi, 1969). This means that when geopolitics before the second world war are discussed, most often the currently used term applied political geography is referred to. It has a stronger link to geography. In this thesis, the term classical geopolitics will be used.

Where Morgenthau and Kjéllen presented a realist view on geopolitics, Wittfogel – a leading intellectual in the German Communist Party – posed that geopolitics represents an bourgeois ideology. Geographical factors, such as location, race, soil, and climate did not directly influence political life, but instead helped to determine politics in societies (O'Tuathail, 2005).

Later, the term 'geopolitics' changed as it was more broadly used, Not only by geographers, but also became very popular among international relations and military experts, diplomats, and journalists (Mamadouh, 1998).

Geopolitical analysis is much more complex than studying geographical aspects, due to globalisation, and the complex interactions between different states (Cohen, 1964).

Al-Rodhan (2014) argues that we cannot forget about the geographical factors. For him, geography includes both features that are fixed as well as structures of human and political geography, such as country size, national boundaries, and historical, religious, and cultural physical sites. The relevance of geography and territory appears sometimes falsely outdated in the current globalised digital world. However, this digitalisation has not surpassed the geographical and cultural historical boundaries. Al Rhodan's view on human and political geography seems closely related to culture in soft power (discussed in 4.2).

Later, geopolitics was defined as a spatial requirement of a state, being a dynamic concept. It investigates the state primarily in relation to its environment and thus provides guidance for practical politics (Höhn, 2011). "Geopolitics is concerned with the spatial requirements of a state while political geography examines only its space conditions. In putting geography at the service of space-conscious politics, geopolitics devotes itself to question of future. It is a discipline that weights and evaluates a given situation and by its conclusion seeks to guide practical politics." (Adhikari, 2013).

Fraser (2009) goes one step further, by posing that globalisation and the impact of the web goes beyond political mobilisation within a country and digital diplomacy between them. He calls this geopolitics 2.0, where there are three significant shifts: (1) states to individuals; (2) real-world to virtual mobilisation and power; and (3) old media to new media.

Although most definitions stay closer to political geography, there is a large shift to a more dynamic and political term. I will use – in line with Haushofer – as my definition for geopolitics: Geopolitics is equal to international politics. This for the following reasons: (1) geopolitics focusses on national power. Although national power is hard to define exactly, it is obvious to be influenced by politics (2) journalists and diplomats use geopolitics more commonly in line with the broader political definition. Using a definition in line with the definition for a broader public makes it more accessible.

National Power

From childhood on we learn to view our social world in terms of who is better, smarter, or favourable above others. Even adults are quick in judging on status through cars, big houses, and career titles (Koski, Xie, & Olson, 2015). Our relative position compared to others is essential for defining social roles and successful interaction (Halevy, Y. Chou, & D. Galinsky, 2011; Savin-Williams, 1979). States, as identified earlier, are just like people in search for their place in the social structure of the world to increase successful interaction with other states. Hierarchy refers to the ranking of members in social groups based on the power, influence, or dominance they exhibit. Status, however, can be measured through social opinion or reputation (Gould, 2002). The terms "rank" and "status" are often used interchangeably or in conjunction as they both represent the position in a social hierarchy (Mazur, 2013).

The position of countries in the social structure of the world is not any different from that in social structures. Power is the ability or right to control people or things and therefore, its measurement is a critical issue. It is one of the main drivers of analysis in international relations and geopolitics, it is an important factor in defining global order (Höhn, 2011). However, for centuries, people have been struggling with its conceptualisation and measurement. It is an important part of cross national analysis, and a successful measurement can assure more precise account of systemic concepts such as 'balance of power', 'polarity', and 'power transition'. National power is difficult to measure as there is the distinction between actual and potential power (Chang, 2014). Three approaches of measurement for national power are identified by RAND: (1) control over resources, (2) control over actors, and (3) control over events and outcomes (Hart, 1976).

Nye uses, for the identification of power, the terms: hard power, soft power, and smart power (Nye, Jr., 2011), respectively. However, what national power is and what it exactly consists of is still not clear, until now.

Characteristics of 'super' powers seem to be that they mastered the following seven dimensions of national power: geography, population, economy, resources, military, diplomacy, and national identity (Kennedy, 1987).

Höhn (2011) identifies two key questions in the formulation of power formulas: (1) which key indicators should be used, and (2) how should the weights of those power components be determined[1]. Within international relations there is always a separation between realists and liberals and the perceived weights and importance of hard and soft power. Spykman (1942) stated that "the search for power is not made for the achievement of moral values; moral values are used to facilitate the attainment of power". So, how should power be measured?

2.2. Approaches To Capture Power

In the following sections current measurements of national power are reviewed based on their ability to show power dynamics.

2.2.1. Power Formulas

As discussed earlier power formulas try to measure final power based on a set of variables for a set of countries in order to make them comparable to each other. The goal of power formulas is to shift from theoretical belief to empirical affirmation (Höhn, 2011).

Power formulas can be operational (2.1) or theoretical (2.2). When they are operational actual numerical results are calculated which will most often lead to a ranking. Theoretical formulas serve as a clarification of conceptual thinking. The difference is the use of concrete variables[2].

$$Power = NominalGDP + MilitaryExpenditures \tag{2.1}$$

$$Power = Economy + Military \tag{2.2}$$

One of the better known and more modern power formula comes from Cline (equation 2.3) (Cline, 1975). He uses political geographical components as well as softer (strategy, national will) components. Whenever either side of the multiplier is zero there is no perceived power. This takes a lot of weight away from the geographical components.

$$P = (C + E + M) * (S + W) \tag{2.3}$$

Where:
P = Perceived Power
C = Critical Mass = Population + Territory
E = Economic Capability
M = Military Capability
S = Strategic Purpose
W = National Will

The power formula of Rummel (1976) (equation 2.4), uses in line with Cline's formula (equation 2.3 softer components next to physical components. With the examples of Cline and Rummel it can already be seen that these power formulas including soft components are already much more difficult to understand and quantify than power formulas focussing on only components of applied political geography.

$$P = C * I * W \tag{2.4}$$

Where:
P = Power to achieve goal
C = Capabilities to achieve goal
I = Interest in achieving goal
W = Will power to achieve goal

[1] For a more extensive review of the theoretical foundations of national power and its operationalisation in the context of political science and international relations see Höhn (2011)

[2] For an extensive review of power formulas see Höhn (2011).

In principal, most power formulas show the inclusion of military, economic, and psychological power in which military and economics form the basis of hard power, and psychological power is the predecessor of soft power. However, the change in importance of the elements of national power over time is a real problem (Cholley, 1942; Höhn, 2011). Therefore, it would be useful if in a dynamic model not only weights could be assigned, but also a convergence to a certain weight could be made over time.

2.2.2. Indices and their methods

Current power formulas are most often published as indices, containing a large number of components and variables to examine the level of national power compared to 'classical' formulas.

Examples are: the Index of National Power (INP), the Global Firepower Index (GFI), the Soft Power 30 (SoftPI), the State Power Index (StatePI), the Fragile State Index (FSI), and many more (Global Fire Power, 2018; In.Europa, 2017b; Kim et al., 2012; McClory, 2017; The Fund For Peace, 2018).

All these indices measure in one form or the other national power or parts of national power. However, these indices are all data measurements or expert-judgements about the past. They do not enable experts to identify possible futures, which is important for the analysis of geopolitics and autonomous developments. Next to that, there are a lot of flaws in the measurements of these indices and the outcomes it generates.

The use of rankings and ratings can be powerful tools within attractiveness and persuasion of soft power. The growing use of indices by governments, non-governmental organisations (NGO's) and campaigners make them highly effective tools in decision making (The Economist, 2014a). However, there are many problems with indices. One of them is the filling of unknown data. The Economist (2014a) discusses the Global Slavery Index, a ranking of 160 plus countries based on the prevalence of slavery, broadly defined to include victims of trafficking, forced labour and child brides. It made global headlines which, however, have been heavily criticised. As the filling of data already introduced difficulties using data from a previous year, this index goes one step further. Data gaps were filled by using figures of other countries instead, sometimes from a completely different continent.

In the Economist the flaws of indices were presented in a satirical way. Providing a guide for building one's own index: (1) "Banish pedantry and make life easier for yourself by using whatever figures are of hand, whether they are old, drawn from small or biased samples, or mixed and matched from wildly differing sources.", (2) "Get the presentation right. Leaving your methodology unpublished looks dodgy. Instead, bury a brief but baffling description in an obscure corner of your website, and reserve the home page for celebrity endorsements. Get headlines by hamming up small differences; minor year-on-year moves in the rankings may be statistical noise, but they make great copy.", and above all (3) "Remember that you can choose what to put in your index, you define the problem and dictate the solution." (The Economist, 2014a). Of course there are many indices that are thought extremely well through, however, it explains the ease with which an index can be built and introduce false information.

The influence and power of indices can be great as can, for instance, be seen in the OECD's Programme for International Student Assessment (PISA). This index rates 15-year old students on academic performance. Germany scored in the 2001 ranking worse than expected which led to school reforms and a 4 billion euro federal education support programme (The Economist, 2014b). To overcome the unintentional use of flawed data indices are discussed in great detail where needed.

2.3. The Use of Power

Power can be used in several ways. It has just been discussed how it is captured, not how it can be used. In this section the earlier coined hard power, soft power, and smart power will be further explained. These terms are used to describe: the power of coercion, the power of persuasion and attraction, and its combination in the form of strategy (Nye, Jr., 2011).

Hard power is the influencing of interests or behaviour of other political bodies or states by the use of economic and military might (Bavardage Diplomatique, 2016). Hard power relates to the use of 'carrots' (inducements) and 'sticks' (threats) (Ilgen, 2006; Nye, Jr., 2003).

The ability to coerce depends on several factors that mainly reflect applied political geography. A large army can function as the muscle of a country, it provides security as it deters enemies. But, if attacked, it can be used to defend itself. However, to be able to grow a large army one needs a large population, and to sustain a large population in its basic needs, greater land mass is needed. Coercion does not always have to be through the use of military capacity. A large economic might can coerce through economic inducements in the pursuit of its goal. Hard power has been part of many debates as its proponents are often war theorists whom argue that bad violence should be countered with good violence (Lackey, 2015).

Soft power, in contrast, refers to co-optive power (Nye, Jr., 1990). It is a nation's ability to attract and persuade (Nye, 2004). It arises from the attractiveness of a country's culture, political ideals, and policies (Nye, 2004).

Although soft power was only named in 1990 (Nye, Jr., 1990), it is not new and has existed in another form. As discussed shortly in section 2.2.1 psychological power focussed on the willingness to fight and the morale of the population. Which arose after the first world war where Germany attributed their eventual defeat to the decline in national morale towards the end (Höhn, 2011). Carr (1946) introduced: power over opinion, showing great resemblance to psychological power. Steinmetz (1929) stated that victory comes to those who can withstand the pressure of war the longest. Although both focussing on the softer social aspects. It has to be noted that psychological and political power take the notion of manipulation and propaganda, different from soft powers goal seeking through persuasion and cooperation.

There are several difficulties following the definitions of hard and soft power. Höhn expresses his concern about the term 'soft power' itself, as 'soft' can be interpreted as anything. He writes that power has two sides. Using an army for the means of war is hard power, where using it for UN peacekeeping it resonates with soft power (Höhn, 2011). Similarly Nye identifies that a significant source of soft power for Japan comes from the success of its manufacturing sector (Nye, Jr., 1990). Economy, usually identified as a source for hard power, becomes part of the soft power definition. It is therefore also interesting that the Soft Power 30 index, coined as Nye as 'the clearest picture of soft power today' (further discussed in section 4.2.1) does not contain any economic variables [3] (McClory, 2017).
For hard power the same could be argued. It is unclear where hard power stops as it is supported by education, technological development, and the government.

I identify two possibilities for interpreting the terms. Due to the emphasis in soft power on cooperation and attractiveness the terms take the notion of 'good cop (soft power), bad cop (hard power)' or 'legitimate (soft power) versus illegitimate (hard power). However, as it was identified that hard power can also be used for good doing, that does also mean that soft power can be used for wrong doing where goal seeking is not always based on persuasion, but possibly manipulation.

Due to the overlapping systems of hard and soft power Nye rather defines power in line with outcomes which he calls smart power. Smart power refers to the combination of hard power and soft power (Nye, Jr., 2006). It is the link between the potential and the actual (Rummel, 1976). Some countries are more effective than others in converting their potential power into actual power (Nye, Jr., 1990).

A country's culture (when it is pleasing to others), its values (when they are attractive and consistently practiced), and its policies (when they are seen as inclusive and legitimate) are all value based (Nye, Jr., 2009). An infinite number of ways can be thought of to interpret the different components of soft power raising questions as how to define pleasing, attraction, and legitimacy. Ambiguous definitions are needed to be able to frame them as needed in diplomacy. However, for modelling, clear definitions are preferred as different definitions can lead to very different models and model outcomes.

In the following chapter modelling methodologies are discussed in combination with how it is possible to overcome ambiguity in definitions and uncertainty in relations.

[3] The Soft Power 30 Index consists of six sub-indices: government, global engagement, culture, education, digital, and enterprises (McClory, 2017).

3

Methodology and Background

Our ability to make predictions is very limited. That little understanding we have about the future comes from current conditions and previous experiences (Shell Global, 2018; Young C. Kim, 2013). To overcome this limitation computer simulation models can be used. However, modelling and simulation, or the scientific approach for issues in the fields geopolitics and international relations have not always and everywhere been welcomed with open arms.

Bull (1966) coins several points on criticism on the use of the scientific approach in international relations. Where his main arguments are that the scientific approach does not enable us to deal with the elusive substances, and value judgement that is needed in specific situations, writing that it is at best useful for illustrative analogies due to the oversimplifications being made. However, Bulls work from 1966 after which computer science, and with it, the use and usefulness of complex simulation models has grown exponentially. Bulls concerns are still valid, but not insurmountable anymore.

In policy analysis, quantitative scenario development becomes more and more a standard procedure. However, in the field of international relations there is, quantitatively, despite the growth, still very little done. There is, however, one simulation model in this area that stands out particularly.

International Futures (IFs) built by the Frederick S. Pardee Center for International Futures, is a large-scale integrated global model. Through IFs the developers seek to understand the state of the world and the future we may see, and lets us think about the future we would like to see. IFs shows world dynamics behaviour, facilitates scenario development and policy analysis. IFs as a whole is built out of smaller model structures such as: education, health, economic, population, socio-political, and more. These sub-structures interact dynamically with each other, and are driven by quantitative data on 183 countries (Hughes, 2009).

Although IFs proved itself as a widely used tool in decision making, it has - in its current state - also some limitations (Hughes, 2009). Four main limitations are that: (1) not all sub-models are well developed, (2) (deep) uncertainty analysis is not possible, (3) visualisations of the model outcomes are outdated by which it becomes difficult to understand, and (4) the model documentation lacks in information, which makes understanding and debating certain structures difficult.

IFs provides power calculations and power scenario development as well. It uses 18 variables to calculate power dynamically over time for scenario analysis purposes. Where it let one assign weights to that set of variables [1] (Turner et al., 2017). Although – due to missing explanations in the documentation – not every variable is clear, the broad picture is assumed to be that their calculations are in line with a realists perspec-

[1] the variables are: (1) population, (2) GDPatPPP (GDP at purchasing power), (3) GDP (GDP at market prices), (4) GDPxGDPPCatPPP (economic-technological capability using GDP per capita at purchasing power), (5) GDPxGDPPC (economic-technological capability using GDP per capita at exchange rate), (6) Govt (government size), (7) MilSpend (military spending), (8) ConvPow (conventional military power), (9) NucPow (nuclear power), (10) NucWarHeads (nuclear war heads), (11) EnergyNetExport (net energy exports), (12) TradeTotal (total trade), (13) FDIInflows (foreign direct investment inflows), (14) AidNet, (15) HumCapWkPopHDI , (16) GovRev (government revenues), (17) Innovation R&D (innovation in R&D), and (18) connectivityICTProd (Turner, Neill, Hughes, & Narayan, 2017).

tive of the focus on economic and military power.

3.1. System Dynamics and Exploratory Modelling

In system dynamics (SD) relations, feedback effects, accumulation effects, and delays of complex non-linear systems are modelled using mainly first order differential equations (Forrester, 1968; Pruyt, 2013; Sterman, 2000). This makes systems modelling – compared to other modelling techniques – an effective tool for answering the proposed research objectives.

Through the processing of various runs and the analysis of its outcomes, the modeller gains a global view of the system as a whole. It would not be possible to develop such a view from just mental simulation, especially because the interactions between elements are difficult and too complex to separately keep track of (Nersessian & MacLeod, 2017)

SD assumes that a system's behaviour is mainly caused by its own structure (Pruyt, 2013). For a geopolitical model it can be argued that this holds true. Next to a world view, analysing a model allows the modeller to identify undesirable behaviour, and change the structure in such a way that desirable outcomes are generated (Pruyt, 2013).

Complex real-world issues are often dealt with through the use of SD in systems that are nor fully closed nor entirely open. While it is undesired to leave potentially important elements out, it is also impossible and undesirable to model 'the entire world'.

Social complex models are characterised by their high amount of uncertainties from which most of Bulls points of critique on the scientific approach arose (Bull, 1966). The level of uncertainty in which a geopolitical model operates is deeply uncertain. Deep uncertainty is the situation in which, due to an absence of knowledge or data, it is impossible to study the system and its relationships exactly and accuratly. Which results in difficulties in: determining the proper model (mathematical) to define interactions between system variables, selecting probability distributions, and the valuing of desirable outcomes (Lempert, Popper, & Bankes, 2003; Walker, Lempert, & Kwakkel, 2012).

(Deep) uncertainty can be partly dealt with in dynamic simulation models by the use of exploratory modelling analysis (EMA). EMA makes it possible to analyse models of these complex systems and can provide the modeller and experts with important insights about the system. General characteristics of the geopolitical system are possible to simulate, but not its individual details (Bankes, 1993). SD and exploratory modelling complement each other, by the ability of SD to cope with complex feedback effects and EMA's ability to overcome system uncertainties.

SD modelling is at times described as theory-rich, but data-poor modelling (Pruyt, 2014). However, with the use of decent SD software it is possible to explore the system by creating data-rich models. In a data-rich model datasets are combined with theory about the system in which a more accurate representation of the real system can be simulated. The following section will explain several modelling techniques and data handling possibilities in combination with SD.

3.2. System dynamics Modelling and the use of Data

In SD there can be made a distinction between operational and non-operational modelling. They differ in that an operational model represents the system as closely as possible (a model about military capacity would include vessels, tanks, soldiers, etc.) whereas a non-operational model represents a general view of that system without the detailed mechanisms at play. This is often due to its complexity and the availability of data. In figure 3.1 an often used non-operational model is shown of the economy with in figure 3.2 its behaviour. It is shown that with a constant growth rate GDP growth exponentially. In this thesis Ventana Systems Inc. Vensim is used to built system dynamics models (Ventana Systems Inc., 2011).

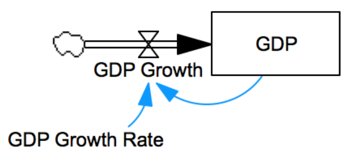


Figure 3.1: Non-operational model of the economy

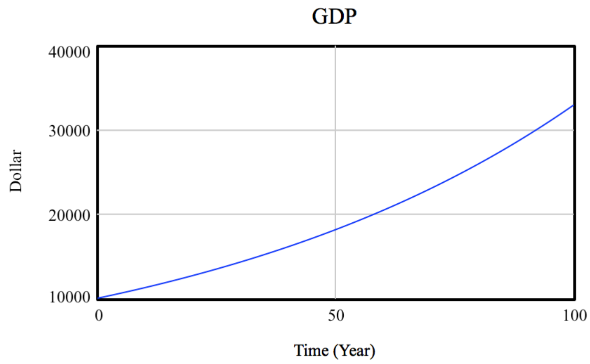


Figure 3.2: Behaviour of non-operational model of economy

However, in this section the population sub-structure is used as an example for building a data-rich operational population model. Figure 3.3 shows a non-operational version of a population model. In this figure the four different variables used in SD are shown as well. These variables are:

1. stock variables, which contain a quantity;
2. flows, able to increase or decrease the stock level;
3. auxiliaries, which are equations of other variables;
4. constants, assumed to be constant over the whole simulation run.

Such a population model is not useful as it does not represent the system that is needed often needed to simulate population dynamics. A main issue here is that it does not take migration into account, but also generalises the population without distinguishing between different age groups and other characteristics. The behaviour shows that the birth rate is lower than the death rate by which the population will decrease. A more operational model is shown in figure 3.5, however, migration has not been taken into account either in this model. The shown model is used as the population sub-model as part of the geopolitical simulation model which will be further discussed in section 4.1.3.

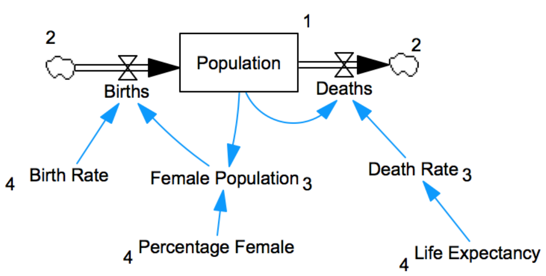


Figure 3.3: Simple population model (not operational)

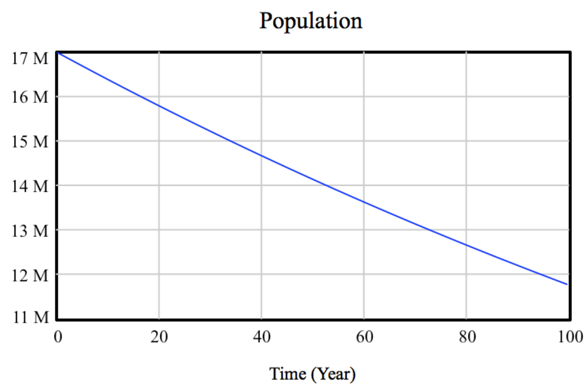


Figure 3.4: Behaviour of non-operational population model of The Netherlands

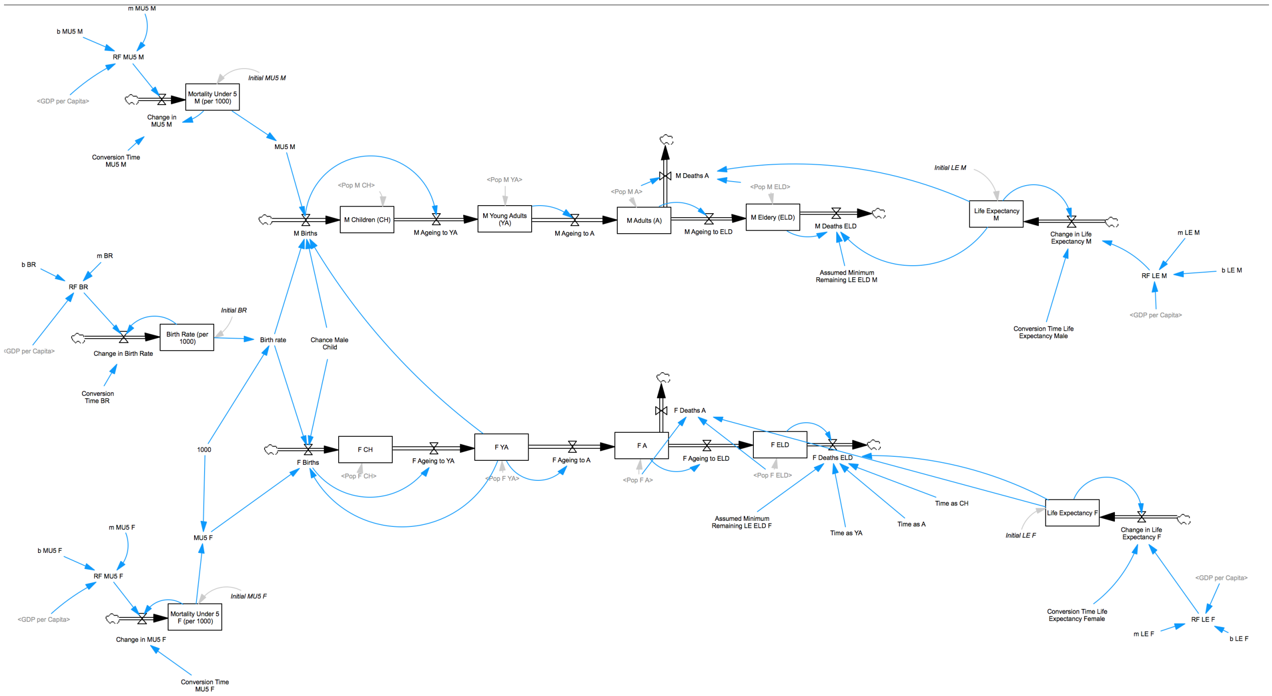


Figure 3.5: Operational data-rich population model using regression curves for value convergence.

Another distinction can be made between data-driven modelling and model-driven modelling. Data-driven modelling is most often used in econometrics where regression techniques are used to derive insights within the available data (Bankes, 1993). Where, in model-driven models the behaviour is driven by the model itself. In SD, data-rich models are possible in which data complements the model structure in producing more representative behaviour.

The IFs model is such a data-rich model. These models are formed by a standard set of algorithms. One of those used algorithms is the convergence process. Which means that a region will converge to a certain manifold. IFs establishes those regression functions by first researching and possibly finding relations within their data (data-driven), they will then use regression analysis to test whether this relationship can be confirmed. After which they integrate this function into their model. The life expectancy at birth (Preston Curve) is taken from the IFs population model to be reevaluated and modelled in Vensim.

Regression Analysis for Variable Convergence

An often used regression function in social science is the Preston curve (Preston, 1975). This curve shows the relation between GDP per capita and life expectancy at birth. Data is well available on GDP per capita and Life Expectancy for most countries. Therefore, the Preston Curve is an excellent fit to test the use of regression analysis and its curve in a SD model.

By analysing data from the World Bank for life expectancy of the year 2016 and GDP per capita for 2010 (for 2010 inflation has been taken into account and is therefore a better data set over all countries) the expected logarithmic relation is found.

Figure 3.6: Relation between life expectancy and GDPPCP (2010)

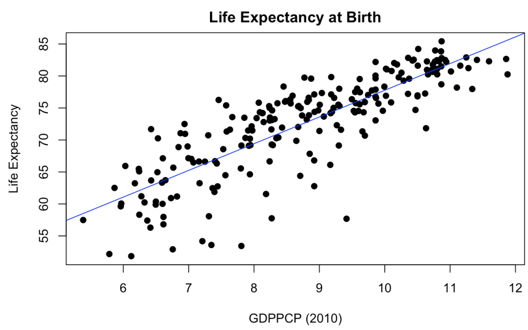


Figure 3.7: Linear regression of life expectancy (logarithmic scale)

```
Residuals:
     Min       1Q   Median       3Q      Max
-17.6518  -2.1431   0.5098   2.9883   9.0911

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  36.0139     1.7157   20.99   <2e-16 ***
x1            4.1759     0.1932   21.62   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.305 on 215 degrees of freedom
Multiple R-squared:  0.6849,    Adjusted R-squared:  0.6835
F-statistic: 467.4 on 1 and 215 DF,  p-value: < 2.2e-16
```

Figure 3.8: Outcomes of linear regression for life expectancy

Using R language to perform linear regression analysis on the relationship in figure 3.7 heteroskedacity is found in 3.7, which means that the errors are not identically and independently distributed. It shows a narrowing noise reaching the top right corner. As heteroskedacity does not introduce bias in the estimates of the regression coefficients it does introduce a bias in the variance estimates of the ordinary least square (ols). As it does not affect the coefficients and the relation is known to be true there is no need to perform any data computation to get rid of heteroskedacity within the data (Johnston, 1963). No further issues were found recreating and fitting the Preston Curve to the new data.

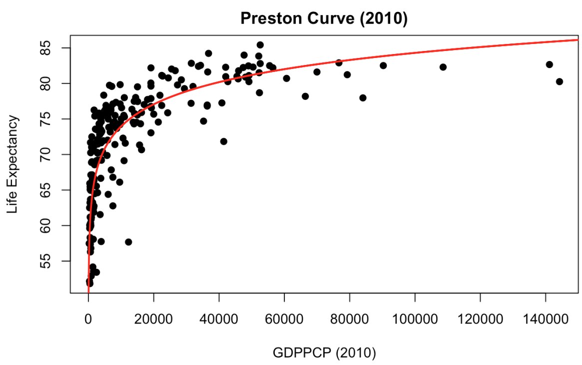


Figure 3.9: Preston curve (2010) on World Bank data

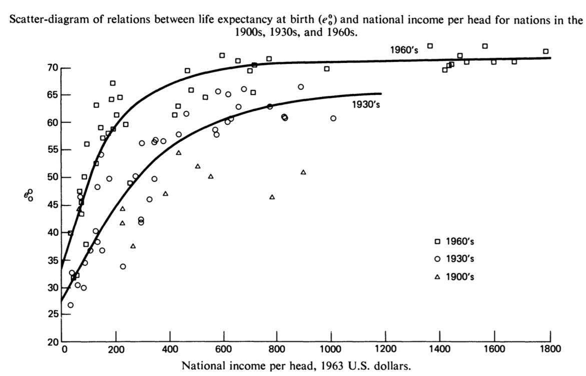


Figure 3.10: Preston Curve 1963 (Preston, 1975)

In Vensim regression functions can be modelled as shown in figure 3.11 in which life expectancy is a stock and a single flow modelled as both an in- and outflow. In which a conversion time is set in the value of life expectancy converges to the value in the regression function.

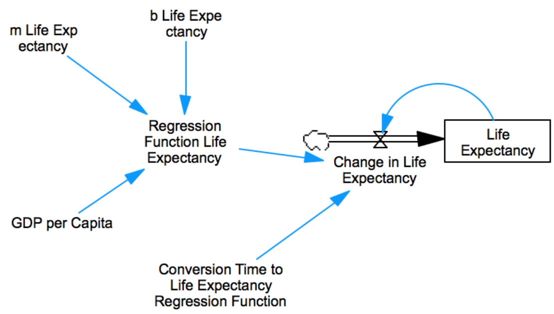


Figure 3.11: Regression structure as used in Vensim

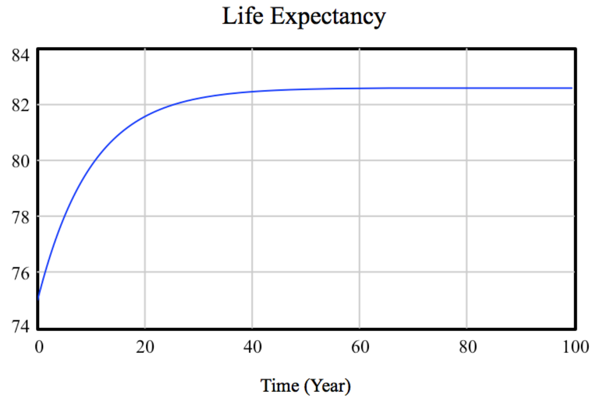


Figure 3.12: Output regression life expectancy Vensim. GDP assumed with a positive growth rate. M = 4.1759 and B = 36.0139.

As not always linear regression is used, but multivariate regression as well overfitting becomes a real problem. It is easy to bring any multivariate correlation to a perfect score of 1.00. However, this without increasing the predictive value. Overfitting is the improvement of description over prediction. This is something that has to be taken into account while in the search for correlations in data, as it is very difficult to find the tipping point of the predictive value. This thesis, due to the complexity of multivariate regression, only uses linear regression.

3.3. Multi-Scale Modelling

As this thesis seeks to simulate shifts in national power it is wanted to be able to simulate for more than one country. It is therefore chosen to use multi-scale modelling at the country level for the geopolitical simulation model discussed in chapter 4. This is done through the use of subscripting in Vensim with which a single scale model can be used to simulate behaviour of other entities of that same scale (countries in this case), making it a multi-scale model.

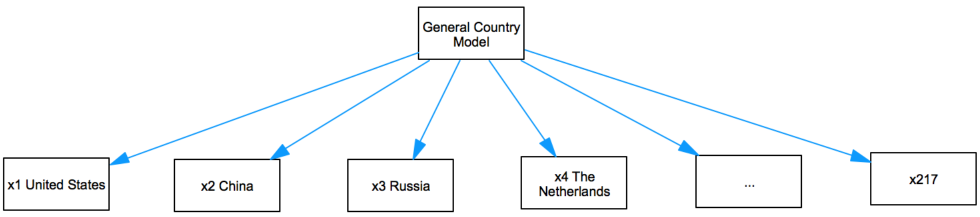


Figure 3.13: Multi-scale modelling for 217 countries

To conclude, the possibilities in the use of data in combination with SD and with multi-scale modelling are significant. Variable uncertainties in the system can be overcome by the use of EMA, and in places where data is available it can support the model for more precession in sometimes complex structures.

In the following chapter, we will return to the measurement and modelling of power

4

Modelling Power

As discussed in section 2.2.1 there is large amount of formulas measuring national power. The components of these formulas often represent a general idea and its impact on the sum of national power. However, to be able to simulate national power over time, one needs to know, besides which key performance indicators the formula consists of, which level of aggregation is needed for the sub-models that, together, form the complete simulation model.

This chapter focuses on the modelling of a geopolitical system dynamics simulation model with as main key performance indicator 'state power'. Sub-models and the difficulties that arise in modelling them are explained and discussed.

4.1. Modelling Hard Power

The following section discusses the components of applied political geography as these fit best with the notion of coercion of hard power. The components are, in order of how much they need the other components: geography, natural-resources, population, economy/wealth, and military power. Uncertainties are portrayed in an orange box in the figures. A full list of uncertainties can be found in chapter 5.

4.1.1. Geography

Geography or the space of a country is an important factor in geopolitics. It is particularly important for the opportunities a countries economy will get. A larger space – most often – correlates with more natural resources, which can not only stimulate the economy in a more direct way, but will also lead to being more self-sufficient.

Table 4.1 shows a summary by Höhn of the perception about the important variables for the use of space in 1998. Höhn shows more than just plain geographical factors in this table. This shows that geography not only consists of geographical constants such as surface area, but also includes parts of infrastructure which is important for the transportation ability in war times.

Table 4.1: Spatial Variables and their Correlation to Power Perception in 1998 (Höhn, 2011)

Spatial variables	**r**	**Data Source**
Highways, Paved (km)	0.790	CIA 1990-2008
Highways, Total (km)	0.708	CIA 1990-2008
Surface Area (km^2) x Per Capita GDP (PPP) [Krahmann Equivalent]	0.544	CIA 1990-2008
Total Internal Area (km^2) + Exclusive Economic Zone (km^2)	0.416	Wikipedia 2010
Agricultural Land (km^2)	0.388	CiA 1990-2008
Surface Area (km^2)	0.258	CiA 1990-2008

In the sub-model of geography data from the CIA World Factbook has been used for the initialisation of three constant variables: land mass, as this influences the possibility of natural resources and the population density. Length of coastline, as this might affect military strategy in the need for marine vessels, or for the possibility for fishery, and the mean level of elevation to portray the amount of mountainous terrain of a country[1][2][3]. The influences of these factors will be discussed in the parts where they influence other variables.

4.1.2. Resources

In hard power, resources can be everything in the pursuet of reaching an objective. However, non-renewable natural resources are particularly important. Oil, Gas, and Coal provide energy, and if sold can boost the economy in times of need. Having many non-renewable resources is a great advantage as it is a large part of being self-sufficient. A country with little non-renewable natural resources always depends on someone or something else.

Within the simulation model only oil resources have been taken into account. The model used is built by dr. E. (Erik) Pruyt. The initialisation of the oil sub-model has been based on data from the CIA WorldFactbook[4]. Within this data set there is a data gap of 40% between imports and exports over the total barrels in the system. It is assumed that this is due to measurement differences between countries. Building a model, these measurement inconsistencies cause problems as such flawed data will initialise a model wrongly from the start.

In this sub-model this has been overcome by the use of extra stocks and flows filling the missing gaps between imports and exports. Due to the complexity of the resource sub-model and its interaction with the system in combination with the current data flaws it is not connected to other sub models. However, the oil price, forecasted based on historical data is used to impact GDP growth. A higher price decreases the GDP growth as oil will cost more. For exporting countries it would be the other way around, a lower oil price will increase GDP. However, it is assumed that in most large exporting countries the growth of GDP ends up in the hands of the elite thus the GDP per capita does not grow. As the simulation model is highly dependent on GDP and GDP per capita (used in all the regression functions) there is chosen to make GDP more dynamic by assuming that: if oil prices rise, GDP declines.

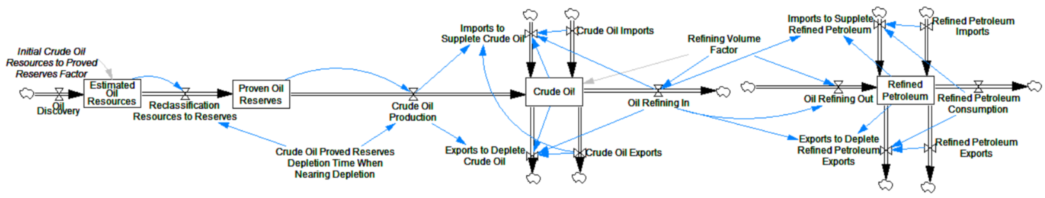


Figure 4.1: Oil sub-model as used in the simulation model (some variables are hidden for a clearer showing)

In future models a wider range of resources should be taken into account. The stock-flow diagram in figure 4.2 shows a simplification of a possible future resource model.

[1]Data for the model initialisation for the land mass has been retrieved from: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2147rank.html
[2]Data for the model initialisation for the length of the coastline has been retrieved from: https://www.cia.gov/library/publications/the-world-factbook/fields/2060.html
[3]Data for the model initialisation of the mean level of elevation has been retrieved from: https://www.cia.gov/library/publications/the-world-factbook/fields/2020.html
[4]Data for the initialisation of the simulation model for oil has been retrieved from: https://www.cia.gov/library/publications/the-world-factbook/

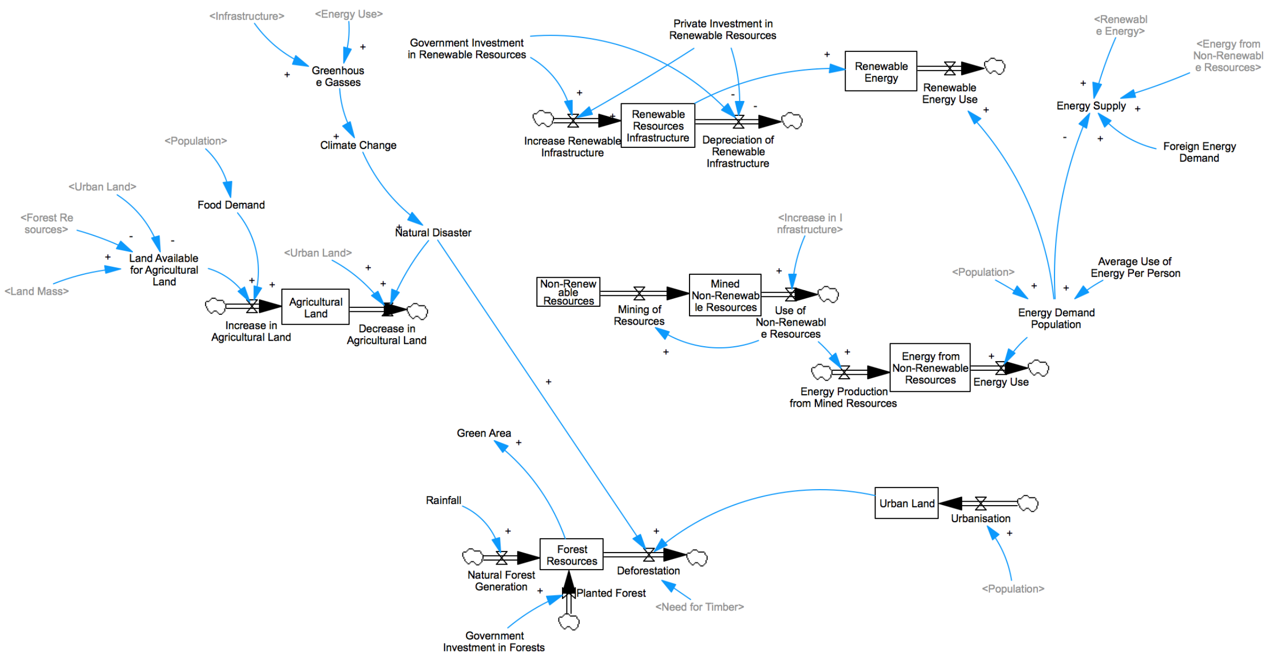


Figure 4.2: Stock-flow structure of possible resources sub-model

4.1.3. Population

Ratzel (1923) stresses that population and geography are equal in importance, but that population figures are better at explaining power. A large space with a small population might in the long run develop into a great power, when the empty land is considered as equaling its potential in the capacity to provide for the population.

Population in political geography is more than population in numbers. Where space sets the boundaries for population in not only land mass, but also the supply and demand for food. A large population with little space makes the country dependent on others if the space does not let the country provide food for the population anymore.

Höhn (2011) summarises the perceptions of population in the year 1988 as well (table: 4.2). Interesting is that the perceptions include a lot of 'psychological' measurements. This has not been included in the simulation model as this sub-system only focuses on population growth and decline in quantities of people.

Table 4.2: Demographic Variables and their Correlation to Power Perception in 1998 (Höhn, 2011)

Demographic Variables	r	Data Source
Population x IQ	0.811	CIA 1990-2008
Population x Cognitive Ability	0.810	CIA 1990-2008
Population x Ethnic Homogeneity	0.662	CIA 1990-2008
Population x Religious Diversity	0.634	CIA 1990-2008
Population x Linguistic Homogeneity	0.612	CIA 1990-2008
Population	0.566	CIA 1990-2008

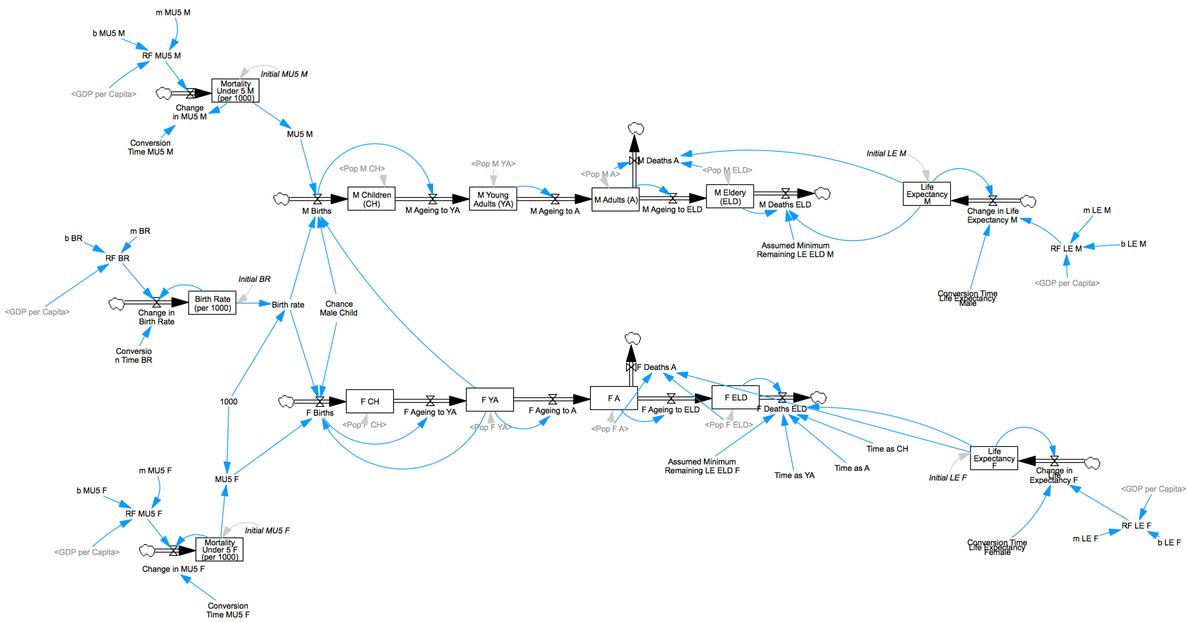


Figure 4.3: Population sub-model as used in the simulation model (some variables are hidden for a clearer showing)

Population has been modelled as two separate ageing chains for men and women. In which four age groups are distinguished: children (0-15), young adults (16-25), adults (26-65), and elderly (65+). Separate dynamics for mortality and life expectancy between male and female come from regression analysis on Data from the CIA World Factbook[5]. All the regression curves are dependent on GDP per capita.

Migration, an important factor in the development of a country's population has not been taken into account in this model as migration and its causes are very complex. Wigman (n.d.) built in his thesis a system dynamics model structure that is able to track a persons nationality. This model could be very useful in the future of modelling migration dynamics as the tracking of people increases the ability to model society and possible societal stress. More on societal sub-model is explained in subsection 4.3.6. However, due to computational limitations it currently only allows tracking for nationality where in the future ethnicity at a regional level could be very useful.

The modelling of the population can be done, compared to other structures, with higher accuracy and more detail. This as there is a useful amount of data available from different sources. Due to the high amount of people in a country the average data on for instance birth rates or life expectancy. Population modelling becomes more difficult in low income countries where there is very little registration of births and deaths. However, rates of regional averages or cultural similar countries can overcome such a data problem.

4.1.4. Economy / Wealth

The economy is very complex and overlaps with many different sub-structures of national power. For the measurement of *Economic Capital* the State Power Index uses three variables: (1) size of GDP according to PPP[6], (2) country rating, (3) number of richest citizens (In.Europa, 2017a). It is unclear what is actually meant with *country rating* and *number of richest citizens*, however, those two variables only count for 20% of economic capital. *Size of GDP according to PPP* accounts for 80% of the economic capital sub-index. Although the State Power Index is used here as example, the use of GDP with such a high weight attached to it is no exception. It is believed that GDP is good at measuring progress this is, however this is not the case.

In the economic sub-model, GDP is the main driver. GDP in term is split into three different production

[5] The scripts for the regression analysis can be found in appendix B

[6] PPP stand for Purchasing Power Parity, converting to PPP makes it able to compare two countries. It is that the exchange rate between two countries is equal to the ratio of the currencies respective purchasing power.

stocks (agriculture, goods, services) based on percentage of GDP data from the CIA World Factbook. Rents from oil have also been taken into account and form, together with the production stocks, the government revenues. Expenditures of a government are initialised by data from the CIA World Factbook, but the expenditures are modelled as either stocks that can chance or - if not needed further in the model - kept as constants. The sub-model is shown in figure 4.4.

Higher expenditures can, for instance, for education lead to a faster convergence to a regression function of the mean years of education. In the case of military expenditures there is an increase in the purchasing of army vehicles.

The economic model is based on a model from dr. E. (Erik) Pruyt, but it has been tried to make GDP dynamic by the use of the changing oil price and a likelihood of declining private investments (if public debt rises private investment is assumed to influence GDP growth negatively) (Bennet, 2017; Lutmer-Paulson, 2017).

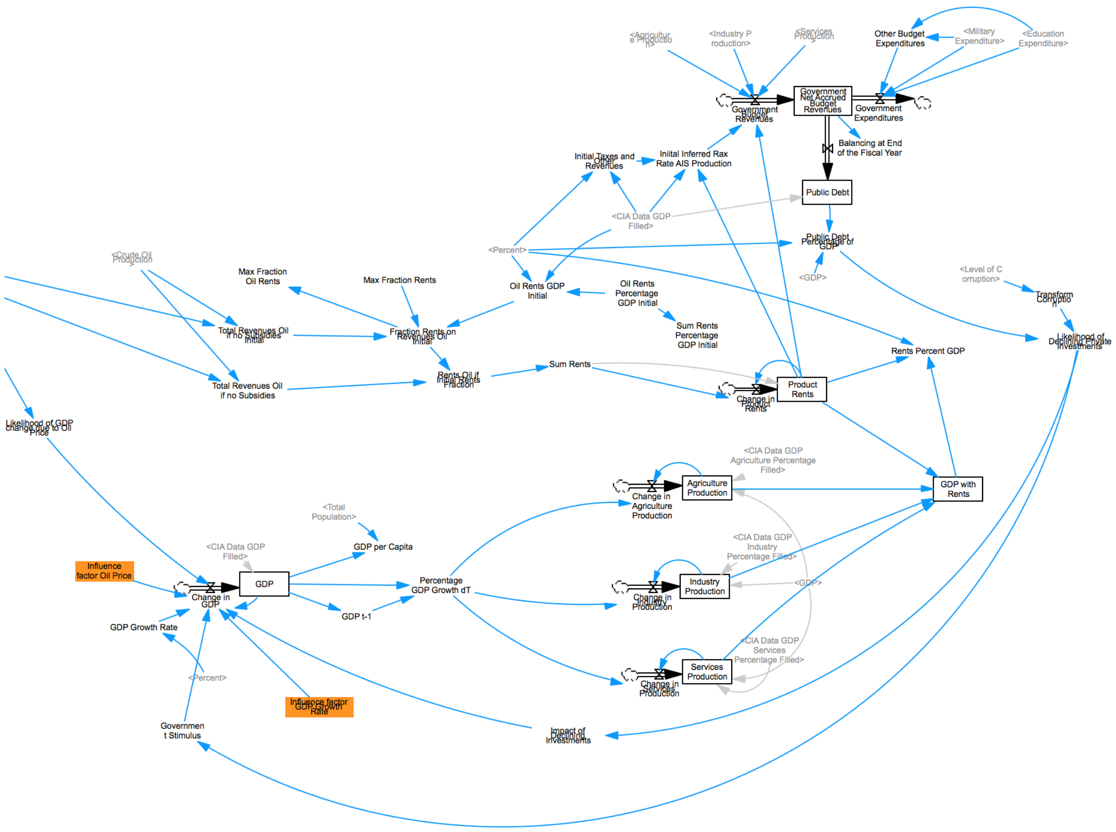


Figure 4.4: Economic sub-model as used in the simulation model (some variables are hidden for a clearer showing)

The use of such a model for simulation (although often used) based on GDP as measurement is not useful as it does not measure what really matters, namely: wealth and the derived power from this wealth.

The use of GDP is not a correct measurement variable for economic capability. Lange et al. (2018); Pilling (2018a) writes "GDP likes pollution, particularly if you have to spend money clearing it up. It likes crime because it is fond of large police forces and repairing broken windows. GDP likes Hurricane Katrina and is quite OK with wars. It likes to measure the buildup to conflict in guns, planes, and warheads, then it likes to count all the effort in reconstructing shattered cities from the smouldering ruins. GDP is good at counting, but a pretty poor judge of quality.". GDP is bad at measuring progress, exactly the thing we think it is good at. Current decision making on the assets of countries has consequences for the opportunities of the population in the future (UNEP & UNU-IHDP, 2014).

Pilling (2018b) wondered why people in Japan seem so well off when their nominal GDP has been stable for the last twenty years. He found that this is partly due to deflation and low population growth, which means that real per capita income is higher than nominal GDP showed. Technology and the quality of services make a difference as well. Pilling writes, "to GDP, an elegant Mitsukoshi department store was the same as a Walmart, and a clapped-out British commuter train did just as well as a Japanese Shinkansen travelling at 200mph and arriving with a punctuality measured in fractions of a second.".

Wealth, created for identifying long-term changes in the well-being of a country is the sum of all assets of that country (Lange et al., 2018). UNEP and UNU-IHDP (2014) write that wealth accounting is especially useful for studies in sustainability issues (autonomous developments) due to its concern with well-being over time.

Without a forward-looking indicator, it is difficult to measure economic process. Wealth – the flow of income that each asset can generate of its lifetime – is by nature concerned with the future. Again, GDP can indicate whether a country's income is growing where wealth is an indicator for the prospects for maintaining that income and growth over the long term. Economic performance may be best measured by evaluating the growth of both wealth and GDP (Lange et al., 2018).

UNEP and UNU-IHDP (2014) suggest that a system approach should be used if it is demanded to perform wealth-based policy analysis. They created a causal link diagram (CLD) for inclusive wealth (figure 4.5). Four feedback loops affect well-being through production and consumption[7].

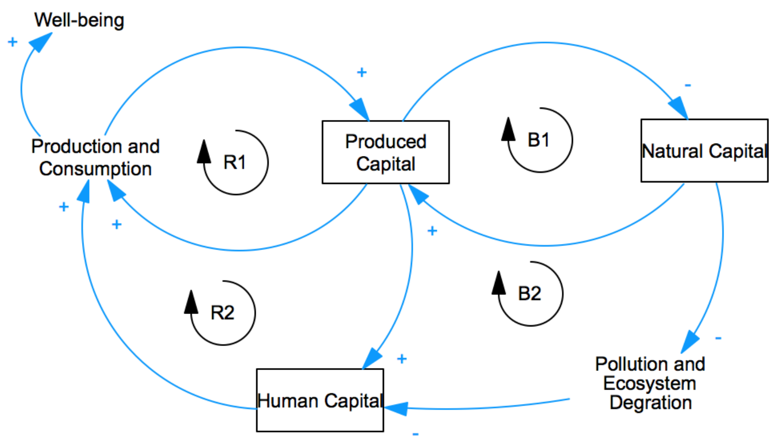


Figure 4.5: Inclusive wealth of an economy as CLD (UNEP & UNU-IHDP, 2014, p.182)

The UNEP and UNU-IHDP (2014) and the Lange et al. (2018) have both published a report on the measurement of wealth. The World Bank, however, just released (2018) a new report. For the following model (see figure: 4.6) the method of The World Bank is used as this reflects the most recent view and knowledge on the measurement of wealth[8]. The World Bank calculates total wealth by taking the sum of each component of wealth shown in equation 4.1.

$$Total\ Wealth = Natural\ Capital + Produced\ Capital + Human\ Capital + Net\ Foreign\ Assets \tag{4.1}$$

In figure 4.6 the components of wealth as described in appendix A of the World Bank report are modelled.

[7] Causal links are portrayed by an arrow. Where a '+' sign means a positive relation (if variable 1 increases variable 2 increases as well). A '-' sign suggests a negative relation (if variable 1 increases, variable 2 decreases, and vice versa).

[8] Both methods are closely similar to each other

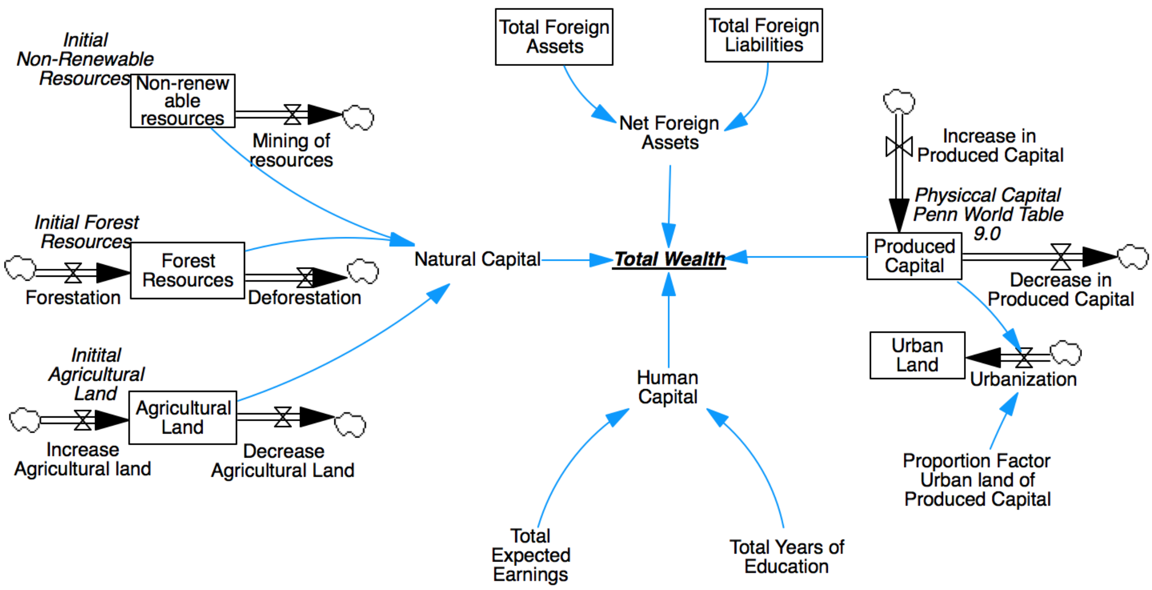


Figure 4.6: Wealth as by (Lange et al., 2018)

There are downsides to the measurement of wealth as well. Few data are available for developing countries, and nations themselves do a poor job keeping count of their own assets. They may even boast about their resources, workforce or infrastructure (Pilling, 2018a; The Economist, 2012).

Another downside is that there is assigned a dollar value on everything, and summing components of wealth implies that nations can lose $100 billion-worth of farmland, gain $100 billion-worth of skills and be no worse off (The Economist, 2012). However, this is only the case when summing wealth. If decent stocks have been taken into account these stocks should be used to compare countries instead of the total sum.

Wealth, as it tries to measure well-being is also highly susceptible to value based weighing of assets. Where Pilling suggests that the well-being of the Japanese is higher due to the level of technology. A neo-luddite might think differently about this topic[9]. In addition is the calculation of wealth based on social or shadow prices, these prices are based on the willingness to pay principle. Next to that, a discount rate for future potential value is used. This discount rate is highly susceptible to bias as it reflects the value a person assigns to different assets (like the weighing of power components).

Roman and Thiry (2016) identify based on the Inclusive Wealth Index (IWI) that empirically the measurement of wealth comes to sustaining the weighted sum of capital assets as monetary values[10]. They ask the question "what happens when the money value of capital assets increases more than the decrease in capital assets?" will it be quantity or quality? This is a value based problem and has to come from literature research about the tipping points of the use of assets.

The problem that comes from the summing of components is a problem in the form of a static measurement. This can be overcome by the use of a SD model. Here the loss of farmland will cause direct feedback effects in the dynamics in the rest of the model. Next to that, the dynamics can be traced and enable us to provide a better overview of causes and possible unwanted tipping points.

Due to the complexity of the economy, the lacking data issue, and the value based issues that arise from the calculation of wealth this thesis does use GDP as driver as described in the beginning of this section.

4.1.5. Military

In the calculation of military power the same problem arises. Often military expenditure is used for the measurement of military power. The State Power Index of In.Europa (2017b) uses: military expenditure, arms

[9] A neo-luddite is someone who opposes to most forms of modern technology (Jones, 2006).
[10] The IWI is based on the 2012 report of the UN (UNEP & UNU-IHDP, 2012)

production and sales, military expenditure as a percentage of the GDP, the number of uniformed officers, and possession of nuclear weapons. However, there is a structural measurement problem in this list of variables: When measuring military power one wants to measure military capability (what can one do with her army). Military capability consists of two things: (1) military capacity, (2) strategy. As discussed in section 4.4 strategy has not been taken into account. For now, the focus is on military capacity as this is actually what the State Power Index is intending to measure.

Military expenditure is a flow, an expenditure over time. Arms production and sales are, again, a flow of a certain production (inflow) and sales (outflow). Military expenditure as a percentage of the GDP is ,again, a flow of expenditure (this has already been taken into account through military expenditure). The number of uniformed officers is a stock of an number of uniformed officers (has partly been taken into account as monetary value in military expenditure), and possession of nuclear weapons is a dichotomous variable with two possible answers: yes or no.

Table 4.3: Militarisation variables State Power Index with comments (In.Europa, 2017a).

Variable	**Measure**	**Comment**
Military expenditure	Flow (measures an expenditure over time)	Useful measurement, but does not represent military capacity.
Arms production and sales	Inflow (production, measures production over time). Outflow (Sales, measures amount of sales over time)	Production could be important, however, has nothing to do with the current status of military capacity. Arms sales are closer related to economic growth than military capacity.
Military expenditure as a percentage of GDP	Flow (measures a percentage of expenditure over time)	A small country could have the same percentage expenditure as a large country with a huge difference in real expenditure.
Number of uniformed officers	Stock (measures a number of officers at a certain moment in time)	Valuable measurement point, however, is one of many.
Possession of nuclear weapons	Auxilary dichotomous variable (yes or no answer)	If measured as yes or no (as the variable suggests) then it is useless, one nuclear head is not the same as forty nuclear heads.

Consistency in the measurement of variables is key to a proper value estimation of the sub-indices. Stock variables are suited for such measurements as these can either increase or decrease. This way depreciation of goods can be be taken into account and countries can be easily measured relatively to each other. Military expenditure is for the military sub-model what GDP is for the economic sub-model.

From a conversation with a student 'Security Studies' from Leiden University I found that the focus of a countries army comes from economic, political, and geographical goals, next to national ideas. However, some parts of a country's defence focus comes from the role it takes within the system. For instance, NATO identifies two types of countries: countries with a direct obligation to defend their country, and countries that do not have to do that. Those countries can therefore specialise in a certain defence type. Norway is identified as the northern flank of Europe and as Norway is important due to its large oil reserves it has to be able to defend itself for 24 to 48 hours. Other countries, such as The Netherlands does not have that obligation and is therefore specialised - mainly due to historical needs for its colonies - in navy forces. For such information very detailed and specific datasets are needed that should be build by the use of expert knowledge.

Military power, as is seen, comes from a capacity and a strategy. This model focusses on capacity. It includes stocks of soldiers, army vehicles, aircrafts, and navy ships. In the model there are a few assumptions on strategy. A country with a coastline length that is 50% or more of its total border will have its main focus on its navy. Whereas a country with no coastline at all will not purchase any navy ships. A country will always have a focus on soldiers and land vehicles. However, as stated before, these assumptions are something that should be consulted with experts. This for the reason that a countries defence strategy is dependent on more than just geographic factors. In order to, in the end, build a final state power formula there is taken the sum of the different defence pillars in combination with their uncertain influence factors. This sum (the variable Total Military Strength) is used in the variable used in the state power formula (see section: **??**). For the initialisation of the military sub-model data has been used from the Global Fire Power Index(Global Fire Power, 2018). A web scraper has been built to retrieve and sort the data to excel. This script can be found in appendix D. Specifics such as described above have not been taken into account in this model.

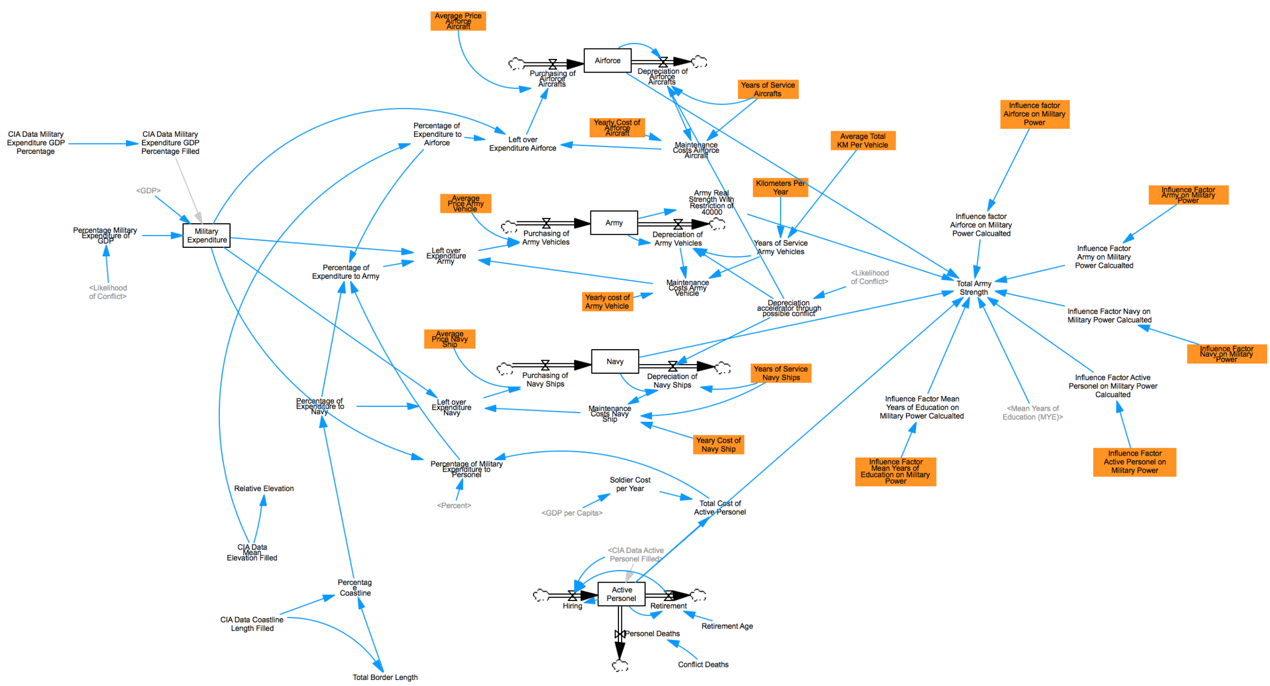


Figure 4.7: Military sub-model as used in the simulation model (some variables are hidden for a clearer showing)

4.2. Modelling Soft Power

Soft power is difficult to measure and quantify as it is such an ambiguous and qualitative term. However, over the past few years there have been several attempts to measure soft power. This resulted in three main indices: the New Persuaders 3, the Rapid-growth markets soft power index, and the Soft Power 30 Index have been established to provide a clearer idea of soft power and its measurement, since persuasion and attraction became more and more valued over hard power solutions (Ernst \& Young, 2012; McClory, 2017, 202).

This section reviews the Soft Power 30 index for the reason that this index has been described by Nye as "the clearest picture of global soft power to date". In this section the modelling structures associated with soft power are discussed as well. For these sub-structures the Soft Power 30 index has formed the basis.

4.2.1. The Soft-Power 30

The Soft Power 30 index, created by the USC Centre of Public Diplomacy, is a comparative assessment of global soft power combining objective data from six sub-indices on: 'government' (measuring political values and public policy outcomes), 'global engagement', 'culture' (measured by the quality of international penetration of its culture, i.e., number of visiting tourists), 'education', 'digital', and 'enterprises'. This is established in combination with international polling for 25 countries (McClory, 2017).

Although coined as the clearest picture of soft power, the index shows the same problems of measurement as were identified in the State Power Index of In.Europa In.Europa (2017b). Since there is little consistency in the variables used for the index, it leads to mis-representative numbers being taken into account. The main concern lies within the use of growth rates which are often measured in percentages. In the sub-index 'Enterprise' the Soft Power 30 uses 'Unemployment rate as a percentage of labour force'. As an example: Although a 2% increase in unemployment in both the US and Denmark relatively show the same growth, the real, absolute outcome or burden which impacts power is not represented in this variable.

The second problem: existing indices, are used as components of the sub-indices in the Soft Power 30 index. This can cause false results and measurements besides the possibility of taken variables double into account. A few indices used are the: Human Development Index score, Freedom House Index score, Gender Equality Index score, Economist Democracy Index score, Press Freedom Index score, and many more (Freedom House, 2018; Kekic, 2007; United Nations, 2018c, 2018d; without borders, 2018). However, even with those well established indices the methodology is not always clear and using an index as a component for another index might lead to, which I would call, 'index inbreeding'.

Also, it is not always clear in what the components are measuring. For instance, "Facebook engagement score for heads of state or government" from the sub-index 'Digital': Engagement could be two things: (1) the sum of all the likes, shares, and comments, or (2) the engagement rate, which is calculated as "(Engagement post(likes, comments, shares) / page likes or followers)*100" (Smitha, 2013).

Besides the confusion on which measurement is used in each interpretation, none of them can correctly measure attractiveness, except if every comment would be analysed. The engagement rate means that someone who has 100 followers on Facebook and gets 1 comment has a 1% engagement. This would, however, be the same for someone with a million followers and 10000 comments. Still, it is a large difference in impact. For both calculations it is not clear if the comments are positive or negative. The likes given on Facebook, do not necessarily have to represent 'likes' anymore (it is possible to choose from different set of emoji's that represent your feeling about that post). For instance, the engagement rate of posts of President Donald Trump can be extremely high. The question is how many comments are negative and positive. It seems there is no distinction in the measurement of engagement. Nye, Jr. (2011) writes that attraction is a complex phenomenon it refers to getting attention, either positive of negative. Nye seems to argue that there is no such thing as 'bad publicity', however, I do believe that bad publicity is not in the interest of soft power. A final main issue with this variable is that it could very easily be boosted by a country to improve its ranking in the index. It is not a robust measurement for power.

Another example comes from the sub-index 'Enterprise' and concerns the number of 'Global patents filed (percentage of GDP)'. First, it is not understood how 'global patents filed' can be a percentage of GDP as certain quantity of something cannot be part of a percentage of something else. Besides that, as discussed in 4.1.4 it is not useful to measure a flow, as in this example, GDP is, since the depreciation of patents is not taken into account. A country could, in that case, grow in power within the metric of global patents filed even though more patents would be depreciated than filed. Besides the number of patents filed (intending to measure the growth of a countries innovation), most often a very particular patent from history will cause cultural soft power growth. The valuation of patents is very uncertain, definitely in relation to soft power gain. And, just like engagement scores, patents files can easily be boosted.

Nye, Jr. (2011) writes that psychologists identified that we are attracted to those who are similar to us, or with whom we identify in a group. People also like similarities in psychical characteristics and shared attitudes. It is here where the Soft Power Index misses an important analysis of soft power: network analysis. It provides a ranking, however, it does not identify which countries are more likely to have power over others due to shared values. They do not discuss how a larger difference in appearance might lead to power growth or power decline.

Lastly, one of the main disadvantages of the Soft Power 30 index is that it only measures 25 countries, which, globally taken, is very little. The variables used are almost impossible to fill in for countries providing little information (like China), or developing countries. One of the purposes of indices is to make a ranking with as goal to make a comparison in performance of various components in different countries. Knowing the score for only 25 countries with components makes the index not very useful. It is understood that the Soft Power 30 index is mostly focused on the methodology of a soft power index than the actual ranking of all countries. However, the methodological part and the choice of variables is, as shown, not sufficient as well.

4.3. Modelling Soft Power

Although the Soft Power 30 is not the best index in its choice for variables, it still provides a guidance and basis for sub-structures in modelling soft power. In the following section the five sub-indexes of the Soft Power 30 Index are discussed and modelled. Full lists of metrics from the Soft Power 30 Index and comments on these metrics their flaws or usefulness can be found in appendix A.1 to A.6.

4.3.1. Culture

Culture, a non-static phenomenon, is the knowledge and values groups pass on through patterns of their behaviour. Culture arises on multiple levels. There are general cultural traits, some are specific and bounded by nationalities, others are particular to small groups and social classes (Nye, Jr., 2011). Nye identifies that more research should be done on the relation between culture and power. He, for instance, asks if the possible attractiveness of Western culture can reduce an increase in extremist behaviour in some Muslim societies?

Power through culture is not always direct, a third party can influence the attraction to a culture. In China, American and Japanese culture proved to be more attractive when shown by South Korea (Nye, Jr., 2011). This relation is shown in figure 4.8 and figure 4.9

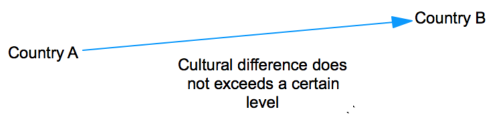


Figure 4.8: Influence of power through culture when cultural difference is limited

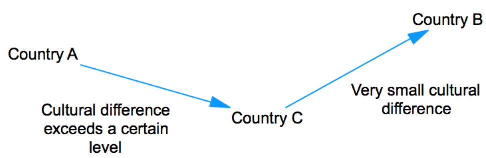


Figure 4.9: Influence of cultural power through a third country when cultural difference is too large

Hofstede, Hofstede, and Minkov (2010) provide a cultural index for 76 countries over six individual dimension: (1) power distance (the degree to which the less powerful members of a society accept and expect that power is distributed unequally), (2) individualism versus collectivism (I versus we), (3) masculinity versus femininity (the preference of achievement and material reward versus cooperation), (4) uncertainty avoidance (degree to which the members of a society feel uncomfortable with uncertainty and ambiguity), (5) long-term versus short-term orientation, and (6) indulgence versus restraint (free versus restraint gratification).

Hofstede is in his index well aware about the psychological traits of in-group (character traits close to us) and out-group (character traits different from us) which makes this index possibly a very useful tool in the understanding of different cultures and attractiveness. There are, however, two large disadvantages: (1) only a limited amount of countries are taken into account and (2) the data is taken from IBM employees, and are thus from a business setting (for which this index was also intended).

The Hofstede Index also shows why the scientific approach to social problems might be rejected. From personal experience with the Hofstede Index I have found that a large number of students where in-dignified by the generalisations that where 'labeled' on them. I believe this is a general difficulty in social science as there are always outliers that can reject the hypothesis. However, the generalisations about those students whom rejected the generalisations were accepted by the other students about that student. In conclusion: generalisations about social structures might be uncomfortable, but generalisations can be found, even in social structures, that can help in understanding the system.

The Soft Power 30 sub-index culture, however, does not represent culture in any way. It seeks to measure cultural attractiveness, but fails in its ability to quantify it by taking variables for which data was available and that could be connected in any way to culture. Their main purpose for the index is measuring the soft power of countries relative to each other. If done as in the Soft Power 30 index then this means that data gathering should be as consistent as possible and dataset should not have too many gaps. As an example, the data on 'total number of tourist arrivals' has been analysed by the United Nations and the World bank. In my perception they both produce good and valuable data sets. However, only the dataset of the World Bank will be discussed, since it is more clear in its methodology than the United Nations dataset[11][12]. The full list of metrics of the sub-index 'culture' with comments can be found in appendix A.1.

The World bank identifies tourist that travel for touristic purposes outside their usual environment (country) for not longer than 12 months. It uses data from: border statistics (sometimes registered only on airports) and tourism accommodation establishments (sometimes: only hotel stays), and a known limitation is that the data account also for the fact that multiple travels within a short period of time might every-time be registered as new arrival. However, two other big flaws can be thought of: with free travel within the EU, it is extremely difficult to register border statistics, and tourist accommodations as Airbnb or B&B are not generally registered. Although these limitations still make interesting figures, such a dataset has too many measurement consistency errors and is only useful to a very limited extent for the measurement of cultural attractiveness in power.

In the cultural sub-model the masculinity index and the long term versus short term orientation index from Hofstede et al. (2010) are used to portray a possible conflict stimulating reaction. Where a higher masculinity score and a lower long term orientation (thus short term orientated) are assumed to show stimulating behaviour.

4.3.2. Global Engagement

Global engagement in the Soft Power 30 Index represents foreign affairs or diplomacy. As for the variables used in 'Culture' this sub-index measures, except for the 'Environmental Performance Index', only stocks. Stocks that, in my opinion, can indeed represent global engagement or diplomacy.

In the sub-model global engagement has been represented to a very limited extend due to a complexity of the system and unavailability of data. No well developed dataset on the number of embassies was found. Therefore, global engagement has been taken into account as 'economic openness' using the sum of imports and exports as percentage of GDP. A difficulty would in the modelling of global engagement still is. How and when does a country extend its global engagement? Should it also contain the building of alliances or their downfall.

4.3.3. Education

In the simulation model the mean years of education has been taken into account through a regression curve based on the mean years of education and the GDP per capita[13].

In future modelling of education I see fit that the model becomes more operationalised by adding more metrics from the Soft Power 30 sub-index (4.10). The model should contain different levels of education. This for the reason that this provides, in a simulation model, a better understanding about the average years and the level of education, which has been identified as important through researching IFs. Also the ranking of

[11] The United Nations database on total number of tourist arrivals can be found at: http://data.un.org/DocumentData.aspx?q=tourism&id=375

[12] The World Bank database on the total number of tourist arrivals can be found at https://data.worldbank.org/indicator/ST.INT.ARVL

[13] Data on mean years of education retrieved from the CIA World Factbook: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2241rank.html

universities which influences the enrolment on the tertiary level could be accounted for. Trend analysis might be an outcome in forecasting possible shifts from university rankings. In the stock-flow diagram (see figure: 4.10 there has been taken into account that the larger the space of the campus an university has (more opportunities) and if a university has enough professors for all students increases the ranking of that university. An increase in spending on education raises also enrolments into different levels of education. Enrolment of international students have been taken into account based on the university ranking.

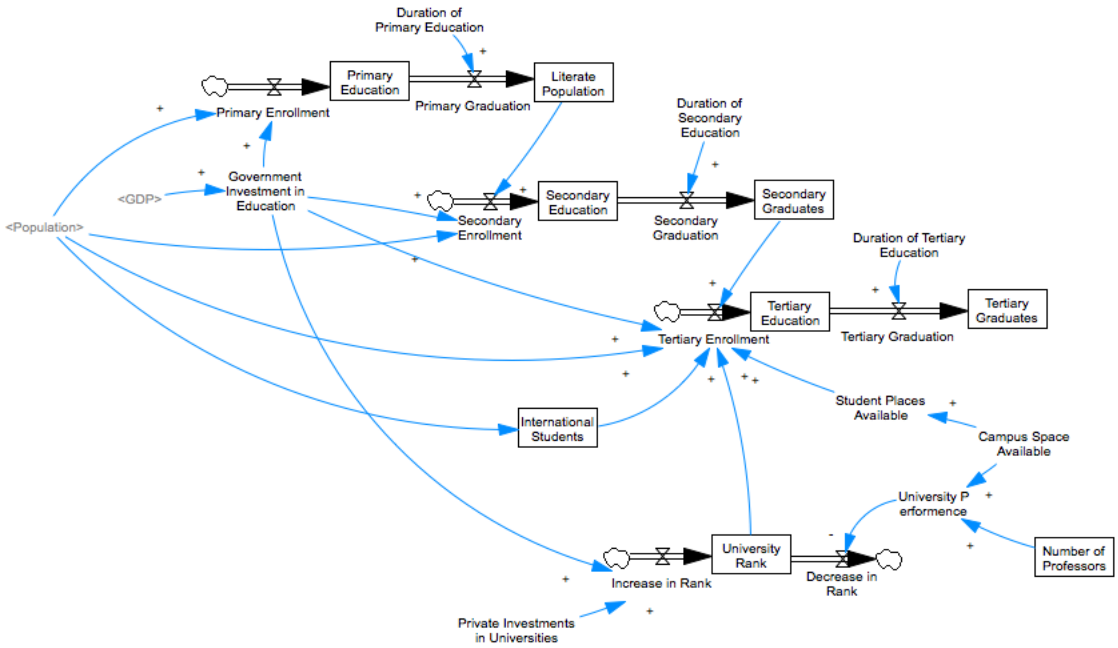


Figure 4.10: Stock-flow structure of education sub-model

4.3.4. Digital

As discussed in the general part about the Soft Power 30 do the components measuring anything with facebook are flawed measurements of the digital sub-index. However, the number of internet users is important for the possible reach of news or propaganda. The digital sub-model has been taken into account by the percentage of of internet users of the full population.

4.3.5. Enterprises

Enterprises partly represent the technological development and global engagement in the geopolitical system. The sub-index 'enterprises' is, due to the complexity, not taken into account in the model.

4.3.6. Government

The governmental sub-index shows that the governmental system is a complex system, as only three components of this sub-index are not indices themselves. For the reasons as described earlier this causes real problems within the calculation of ranked scores.

This government sub-model as used in the simulation model is the largest and most non-operational sub-model. As I believe that next to governmental factors societal factors are important as well to take into account in the growth or decline of state power this sub-model contains a wider set of variables than is used in the Soft Power 30 Index. The government sub-model consists out of three parts: growth of ethnicities, government, and potential for conflict.

Ethnicities

Within the structure of the growth or decline of ethnicity groups the four largest ethnicities within a country and one combined group for the rest of the ethnicities are used to simulate ethnicities. The growth and decline of the five ethnicities are dependent on an uncertain growth rate. Keen (2008c) writes that in an established ethnicity the rapid growth of other ethnicities might lead to uncertainty about the future of their ethnicity. Based on Keen it is assumed that if the two largest ethnicities become almost evenly sized they influence the emerge of conflict positively.

Vogt et al. (2015) built a dataset family that provides information on "ethnic groups" access to state power, their settlement patterns, links to rebel organisations, transborder ethnic relations, and intra-ethnic cleavages". Although there is a large amount of usable information in this dataset only ethnicity sizes of the four largest ethnicities are used to initialise the ethnicities in the model. The model of the ethnicity sizes is shown is figure 4.11.

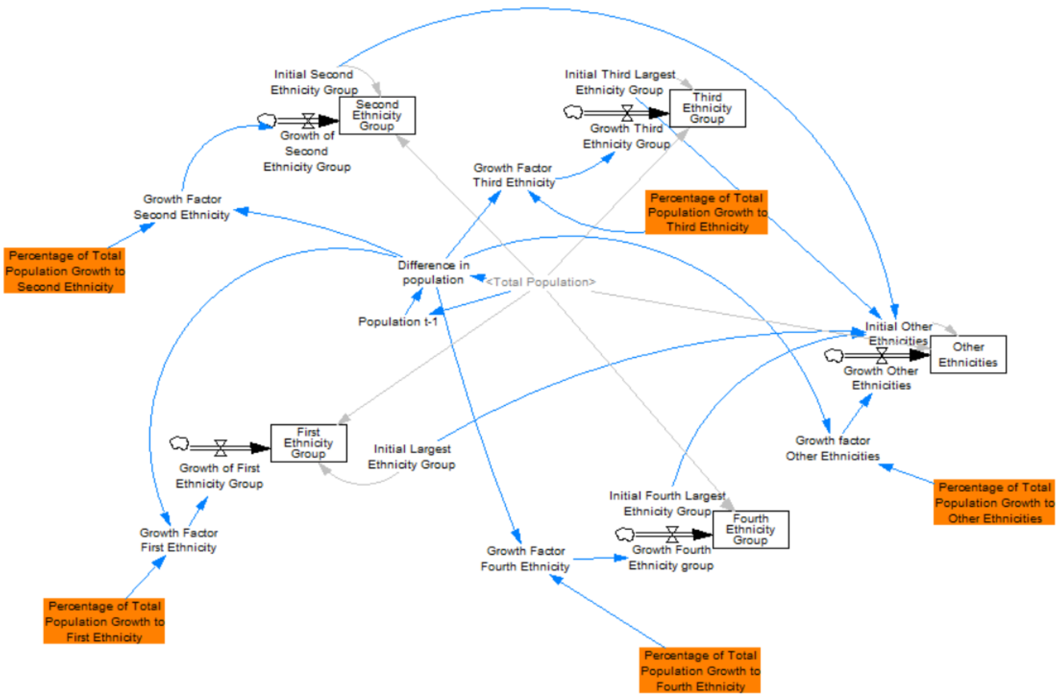


Figure 4.11: Ethnicity sub-model as used in the simulation model (some variables are hidden for a clearer showing)

Government

Within the sub-model of the government there are several non-operational structures that drive the model behaviour. The government is driven by: economic openness, independence of judiciary, corruption, free media, and the level of democracy. Due to their complexity they are modelled as a single stock with a flow that acts as an inflow as well as an outflow.

Economic openness is based on trade as percentage of GDP (Zimmerman, 2017). As its initialisation is World Bank data is used [14]. As economic openness, new trade, new international relations are very complex its growth or decline is identified as an uncertainty.

Economic oppeness is part of the driver of corruption (Quarmyne, William, & Arhenful, 2013). Corruption is driven by both driven by the independence of judiciary and economic openness. It is initialised by the Corruption Perception Index score of 2017 (Transparency International, 2017). Corruption in itself is a driver for the level of free media and the change in the level of democracy. Data gaps have been filled by the use of region average by taking the average of the known values of the countries from a certain region.

[14] Data is retrieved in August 2018 from https://data.worldbank.org/indicator/NE.TRD.GNFS.ZS

As part of the process of authoritarianization it is assumed that when there is decline in the level of democracy there will be decline in the independence of the judiciary as well. However, a higher level of judiciary independence makes it harder to change its independence. The independence of the judiciary is initialised by a sub-index score: 'government powers are effectively limited by the judiciary' from the Rule of Law Index 2018 (World Justice Project, 2018). The process of authoritarianization follows the structure of: deliberately install loyalists in key positions of power (particularly in the judiciary and security services), neutralise the media by buying it, legislate against it, and enforce censorship. Finally, become re-elected Kendall-Taylor and Frantz (2016).

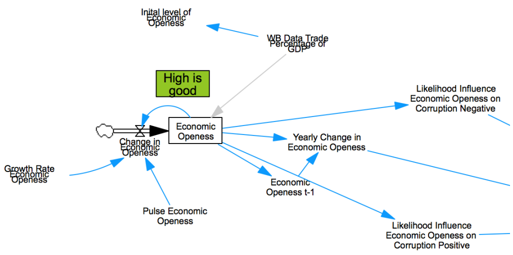


Figure 4.12: Economic openness sub-model as used in the simulation model (some variables are hidden for a clearer showing)

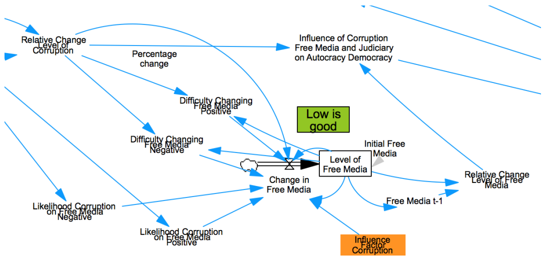


Figure 4.13: Free media sub-model as used in the simulation model (some variables are hidden for a clearer showing)

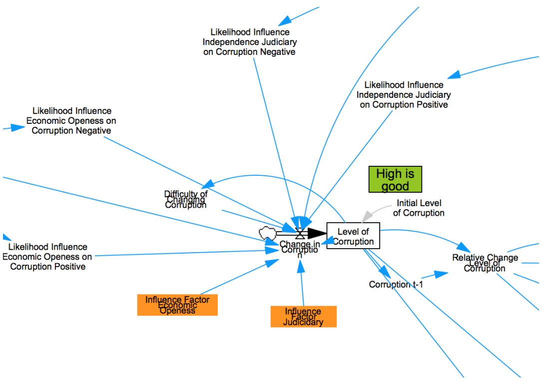


Figure 4.14: Corruption sub-model as used in the simulation model (some variables are hidden for a clearer showing)

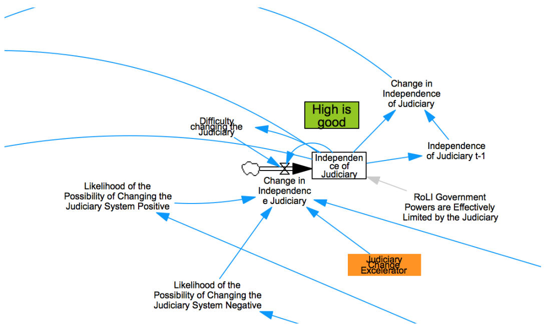


Figure 4.15: Judiciary sub-model as used in the simulation model (some variables are hidden for a clearer showing)

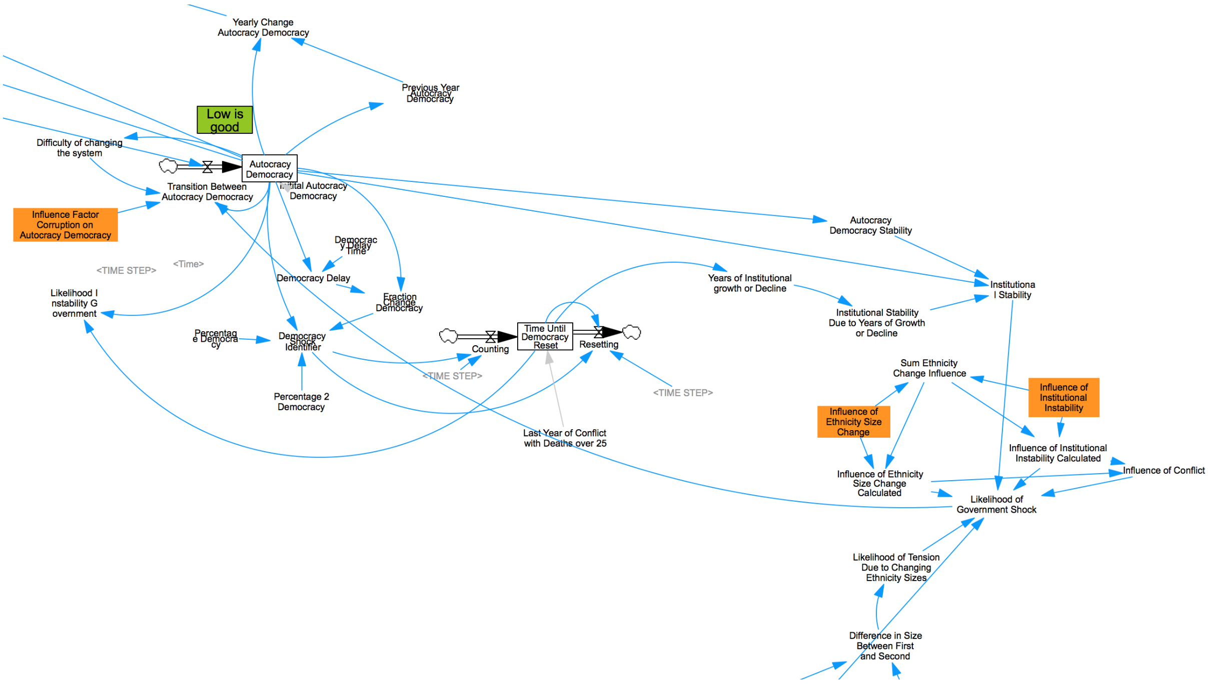


Figure 4.16: Change in level of Democracy sub-model as used in the simulation model (some variables are hidden for a clearer showing)

Conflict

The potential for conflict sub-model has been mainly based on Keen (2008c) in which he engages in the debate of the emergence of conflict. A conventional believe about the emergence of civil war is that grievances create conflict (grievance driven). Collier and Hoeffler (2000) suggest that the control over primary commodity exports (diamonds or timber) is the cause of conflict. The grievances from these conflicts are then used to induce insurgency groups to finance further conflict (greed driven). In 2004 Collier and Hoeffler retract this statement in their revised article stating that grievances are indeed part in the emergence of conflict but has less explanatory power than economic factors (Collier & Hoeffler, 2004). As the change in the level of democracy is already driven by the economic openness and thus economic factors more focus is put on the modelling and simulation of the emerge of conflict through grievances.

Keen (2008a) states that one of the most important roots of violence is a sense of having been humiliated. Which I understand as part of "historical grievances" (using the terms provided by Collier and Hoeffler) from previous conflict. In the simulation model this has been accounted for by using the conflict database of Vogt et al. (2015) in order to calculate the sum of historical civilian deaths as initialisation of historical grievances.

Next, propaganda is identified to have a large influence on the creation of ethnic diversity (Keen, 2008b). This has been the case in Rwanda between the Hutu's and Tutsi's (Keen, 2008a), in Nazi Germany against the Jews, and to a lesser extent the creation of diversity between Sunni's in Saudi Arabia and Shia's in Iran after the Iraq war.

In the simulation model propaganda has been taken into account by the reach of media and impact of discrimination. The likelihood for discrimination is linked to the level of democracy and media freedom. The reach of media is driven by the percentage of internet users which are driven by a regression function of the percentage of paved roads in a country and the GDP per capita in what is assumed to represent the level of infrastructure[15]. Expecting that with a higher percentage of paved roads infrastructure is 'up to date', increasing the amount of internet users. The amount of internet users is then used to calculate the reach of media which can possible be used for propaganda against ethnicity groups.

The discussed Hofstede cultural database has been used to portray a possible conflict stimulating reaction. The sub-indices 'masculinity' and 'long term versus short term orientation' are used in order to form

[15] The data on the percentage of paved roads comes from the CIA Factbook. Retrieved in August 2018 from: https://www.cia.gov/library/publications/the-world-factbook/fields/2085.htmlho

this possible reaction (Hofstede et al., 2010). Masculinity reflects a society with large differences in the way men and women are expected to behave. Boys and men are expected to have a tendency to fight and earn the money to provide for the family. Where girls and women are expected to be the caretakers of the family. In the simulation model a short term orientation or lower score in the index is used to represent a higher likelihood of conflict stimulating reaction. This as it is expected that a reaction in such a culture is more often based on emotions.

The total likelihood of conflict is calculated by the impact of discrimination, current grievances, and the likelihood of a conflict stimulating reaction. As it is uncertain to what extent these factors influence such a likelihood these variables have been assigned influencing variables that will later be used in EMA. The sum of these influencing variables always adds up to one.

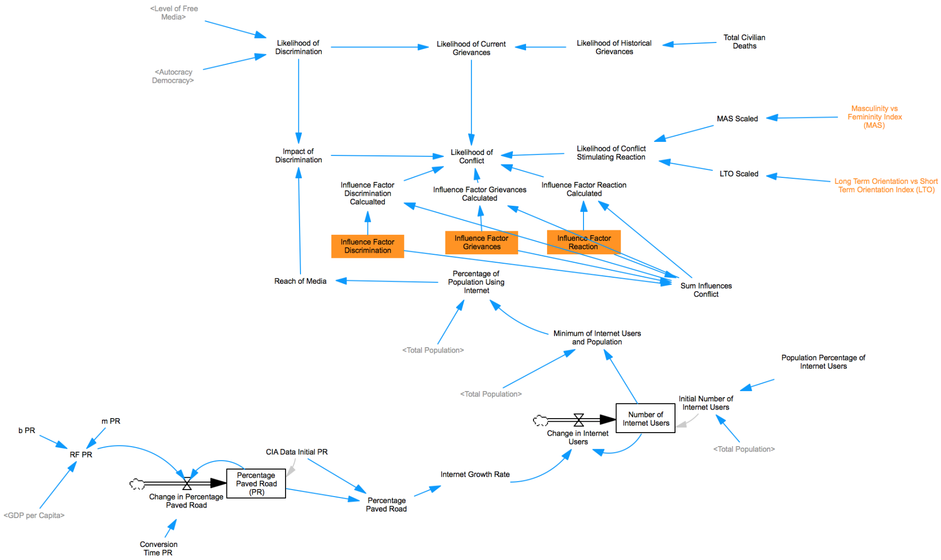


Figure 4.17: Likelihood of conflict model structure.

Finally, some studies found an ‘inverted U’ shape in the level of democracy and the emerge of civil war (Fearon & Laitin, 2003). This has been taken into account by modelling that whenever the level of democracy is very low (autocracy) or very high (democracy) there is more stability. When it swings in the middle (anocracy) then there is a higher chance of a government shock (a rapid change in the level of democracy). The likelihood of conflict also influences the likelihood of a governmental shock.

4.4. Modelling Smart Power

Nye rather defines power in line with outcomes which he calls 'smart power'. Smart power refers to the combination of hard power and soft power (Nye, Jr., 2006). Some countries are more effective than others in converting their potential power into actual power (Nye, Jr., 1990). Höhn (2011) writes that the term replaces labels of old concepts, smart power is the new label for strategy.

Nye does not argues against that, he writes that a smart power strategy answers five questions: (1) What goals or outcomes are preferred? (2) What resources are available and in which context? (3) What are the positions and preferences of the target of influence attempts? (4) Which forms of power behaviour are most likely to succeed? And (5) what is the probability of success? (Nye, Jr., 2011).

Modelling smart power is extremely difficult. This for the reason that smart power is a strategic study that is specific for every situation. Strategy building should be done in co-operation with experts. For this reason smart power is not taken into account in the simulation model.

4.5. Difficulties in Modelling Power

The main difficulty lies in the complexity of the system. An operational model of the economy is almost impossible to build as economists are still figuring out how the economy exactly works as there are also qualitative factors involved. For instance, private investments are largely dependent on the trust that companies have in the stability of a country.

Next, the difficulties grow in the modelling of soft power due to more qualitative relations. Qualitative causal relations are more difficult to establish as data to establish these relations is less available or less consistent. Within the literature the relations found are often not based on statistical analysis, or are of little significance when measured. Often research is also contradictory as other datasets or other techniques are used to establish relations.

Then, due to the high level of uncertainty it is often not clear that a relation is causal or it is a correlation. In system dynamics causal relations are used to model a system. The difficulty in having uncertainty in the relations between variables or systems is that EMA is not able to explore the uncertainties of those relations. However, if researched well, generalisations about the relations between systems can be made over which EMA is able to explore.

In the following chapter exploratory modelling and analysis is performed and the simulation results and difficulties in modelling are discussed to a further extend.

5

Exploratory Modelling and Analysis on Geopolitics

In this chapter, Exploratory Modelling and Analysis is performed on the geopolitical model (discussed in the previous chapter) and is used to identify power shifts in the geopolitical arena. The desired key performance indicators (KPI's) of the model have been identified as state power and the level of democracy, showing shifts between an autocratic regime and a democracy. Throughout history and within the current international power field the United States, China, and Russia have been the large powers. Currently the United States are expected to decline in the power hierarchy and China will rise quickly (Yan, 2006; Zakaria, 2017). For Russia, once a great power, it is uncertain if its power status will rise or will keep declining. Chang (2004) writes that Russia is the number two power in the world, but other research was unsure about Russia's power and expected it to be more closely resemble the power of Germany (Saradzhyan & Abdullaev, 2018). Analysis of the Unites States, China, and Russia has been performed since their power dynamics have been more discussed that that of others. First, the model settings for the EMA Workbench are explained, then the results of the simulations. At the end of this chapter the model in relation to the results will be discussed.

5.1. Setting up the EMA-Workbench

Exploratory Modelling and Analysis uses computational experiments to analyse complex systems while taking uncertainties into account (Bankes, 1993). In this case the system to be analysed is geopolitics. Exploratory modelling and analysis on a system dynamics model is done by using the EMA Workbench (Kwakkel, 2017). In this workbench it is possible to assign uncertainty ranges to a set of variables. In this study, a distinction is made between two types of uncertainties: model uncertainties and state power uncertainties. The differences between those are that the model uncertainties (5.2) influence the model behaviour and the state power uncertainties (table: 5.1) only influence the final state power score.

The state power variable is a power formula with a value relative to the other countries within the model. In this formula mainly variables discussed in section 4.1 hard power are used. The level of democracy is the only variable that portrays a part of soft power. The variables included are: total population, land mass, GDP per capita, army strength, and the level of democracy. An uncertain coefficient for each of the variables described above has been introduced since it is uncertain to what extent variables influence the final state power score. For these uncertain coefficients, a range between 0 and 1 has been assigned, representing a percentage. Within the model it is calculated that the sum of the uncertainties always adds up to one, so that variables will not be overvalued or undervalued in the calculation of state power (see figure: 5.1).

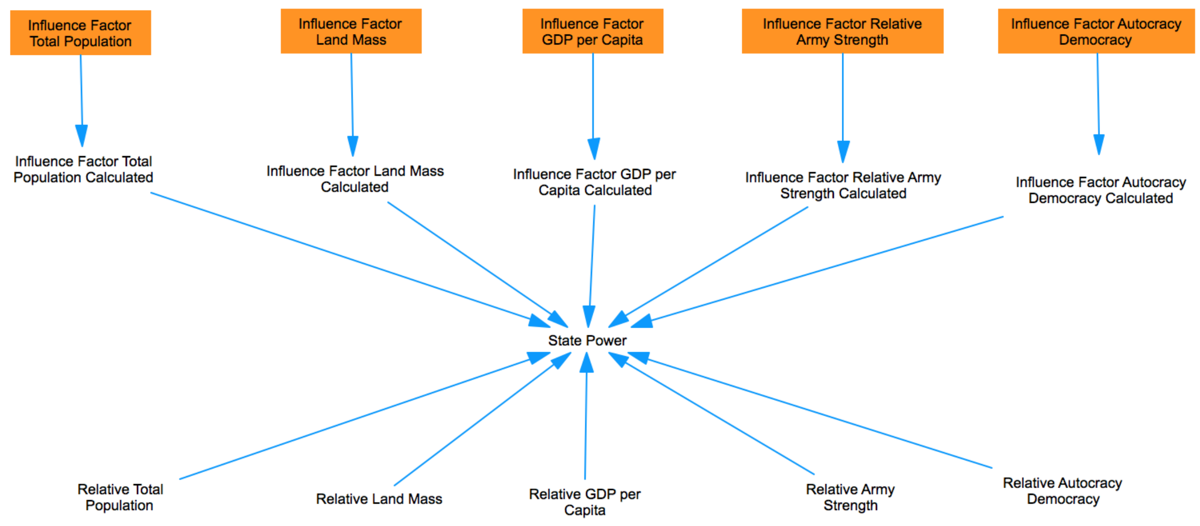


Figure 5.1: Model structure to calculate relative state power.

Table 5.1: State power uncertainties

Uncertainties for Power Sum of Total State Power	**Low**	**High**
Influence Factor Total Population	0	1
Influence Factor Land Mass	0	1
Influence Factor GDP per Capita	0	1
Influence Factor Relative Army Strength	0	1
Influence Factor Autocracy Democracy	0	1

Table 5.2: Model uncertainties

Model Uncertainties	Low	High
Influence Factor Corruption on Autocracy Democracy	0	1
Judiciary Change Excelerator	-0.01	0.01
Influence Factor Corruption	0	1
Influence Factor Judicidary	0	1
Influence Factor Economic Openess	0	1
Influence of Ethnicity Size Change	0	1
Influence of Institutional Instability	0	1
Conflict		
Influence Factor Grievances	0	1
Influence Factor Discrimination	0	1
Influence Factor Reaction	0	1
Ethnicity Growth		
Percentage of Total Population Growth to First Ethnicity	0	1
Percentage of Total Population Growth to Second Ethnicity	0	1
Percentage of Total Population Growth to Third Ethnicity	0	1
Percentage of Total Population Growth to Other Ethnicities	0	1
Percentage of Total Population Growth to Fourth Ethnicity	0	1
Uncertainties for Power Sum of Military		
Influence factor Airforce on Military Power	0	1
Influence Factor Army on Military Power	0	1
Influence Factor Navy on Military Power	0	1
Influence Factor Active Personnel on Military Power	0	1
Influence Factor Mean Years of Education on Military Power	0	1
Resources		
Influence factor Oil Price	0	1

Through the use of the EMA Workbench, 500 different scenarios have been simulated, each initialised as constants with a random value within the set uncertainty ranges over a simulation time of 35 years. The results of these 500 simulation runs with respect to the KPI's 'state power' and 'autocracy democracy' (level of democracy) are discussed in the following section.

5.2. Model Results

In the following sections the model results of the United States, China, and Russia are discussed. It is expected to see a wide variance of behaviour in both the state power variable and in the level of democracy. This because the model has simulated for 500 scenarios initialised with random numbers on the uncertain variables.

5.2.1. The United States

The results shown in figure 5.2 show the change of state power over time of the United States for all runs. It can be noticed that for all runs the development is static. Some of the model scenarios show a growth in power after which it declines again to approximately the initial level of power. But this holds true for almost every run. Most or all countries are static in their behaviour since the state power variable is relative to other countries. In combination with the extremely static model behaviour in the level of democracy shows that there is none to very little dynamic behaviour in the model in certain situations (as in the case of the United States). This does not in any way represent the expected behaviour of decline which was found in the literature. In the model discussion section, the different model behaviours will be discussed on what might drive them.

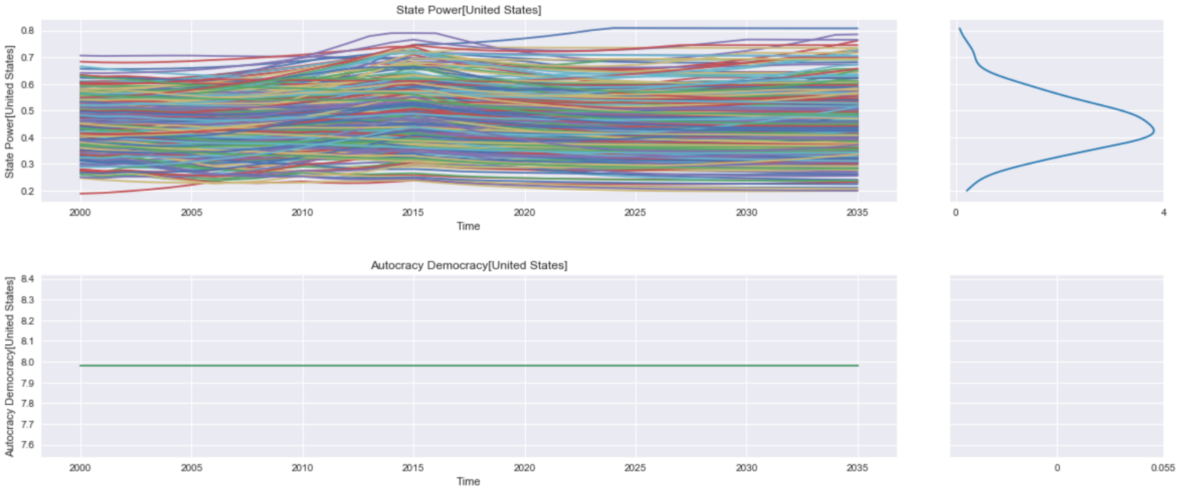


Figure 5.2: Changing power of the United States

5.2.2. China

For China the model shows very static behaviour for state power as well. The level of democracy, however, shows an interesting behaviour. Some of the runs are at first static after which they grow exponentially and then balance out to static behaviour again (as seen as a characteristic S-shape). As for the expected growth of China's national power, it is not seen.

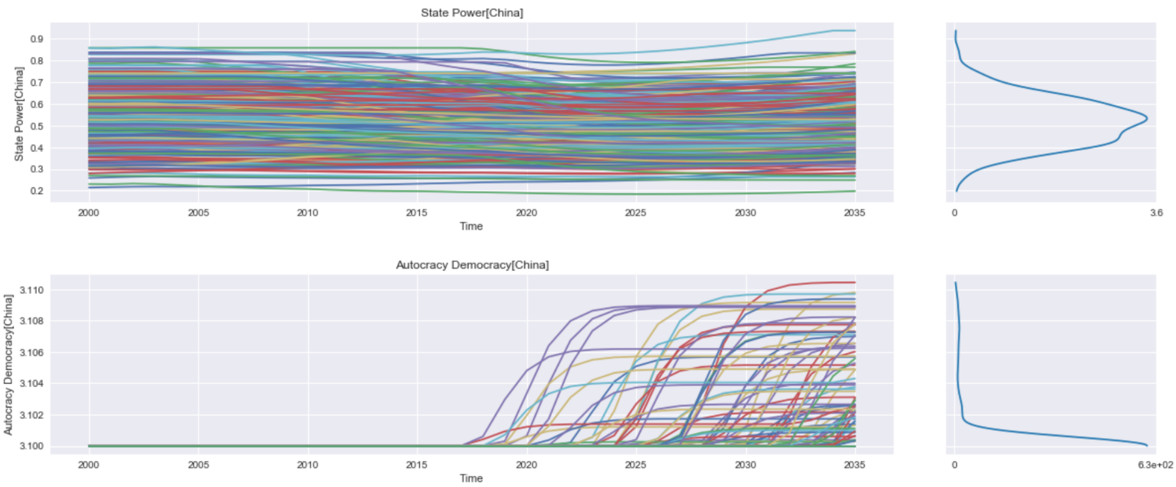


Figure 5.3: Changing power of China

5.2.3. Russia

For Russia, similar behaviour to that of China is seen, although it is even more static. The level of democracy is static in most scenarios as well. Only in a few the S-shaped behaviour occurs.

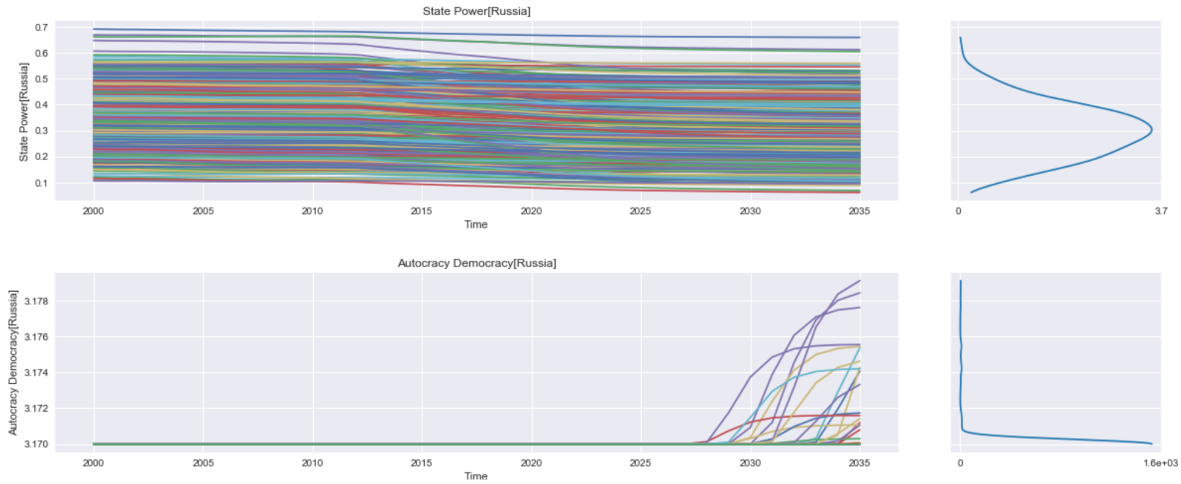


Figure 5.4: Changing power of Russia

5.3. Model Discussion

The model outcomes show very static behaviour for all three simulated countries, meaning that the model is less dynamic than was expected. Several possibilities may have caused this static behaviour, some of which are:

1. The real world system is more static than expected,
2. Oversimplification of the modelled sub-systems,
3. Lacking knowledge about the relations between systems,
4. Modelling choices,
5. Missing or lacking data.

Each of these possible causes will now be further discussed.

5.3.1. The real world system is more static than expected

It might be that the real world system is more static than expected as well. A lot of changes might occur in 35 years, but in a stable democracy during 25 years not many changes are expected. I do believe that the real system is not static. This as all 500 simulation runs show static behaviour. This in combination with the S-shaped behaviour in China's and Russia's level of democracy. Therefore, I assume that the static behaviour is driven by the model.

5.3.2. Oversimplification of the Sub-Systems

This problem occurs mainly in the government and societal model where extremely simplified model structures are used to represent a full sub-system. Feedback effects within the system are not modelled and feedback effects between the systems are difficult to establish. Due to the way the sub-systems are represented in the model (by a single stock and a single flow) it is very easy to generate exponential growth or decline when balancing effects between the systems are not well defined or unknown. In such cases, more knowledge is needed about the system to operationalise it. However, for operationalisation, data is needed that is often missing. For instance, very little is known about corruption, but it plays an important role in the government and society. Additionally, data on corruption is especially scarce for countries where corruption is assumed to be a more dominant factor. Finally, as the stocks represent an index value instead of a real world quantity it is more difficult to define the specific relationships between the systems as these index values already represent a simplified system on their own.

5.3.3. Lacking Knowledge About the Relations Between Systems

Causal relations and feedback effects are highly dependent on literature and expert knowledge in such 'qualitative' systems as the government or society. However, these relations are often not statistically proven, or only marginally significant. The lack in the ability to establish clear relations between systems makes this system deeply uncertain. EMA is not able to overcome the uncertainties in the relations between systems. It can only handle uncertainties in the distributions, quantities, or frequencies of uncertain variables.

Next, as already discussed in the last chapter, if the relations are uncertain there is a high probability that instead of a causation a correlation is modelled as a causation. This could cause static behaviour as there might be a missing third variable that actually drives the dynamic behaviour.

Difficulties and uncertainties in establishing causal relations in multi-scale models are also driven by the fact that model behaviour is often driven by very specific events. One example is the idolatry of the Tutsi's by European anthropologists leading to a civil war in Rwanda (Keen, 2008c). It was expected that a geopolitical SD model would not be able to simulate such specific dynamics, but that it could simulate a general idea about the system. This does seem the case analysing the model. For instance, the relation between economic openness and corruption is very uncertain as it is unknown what exactly affects certain corruption variables (Quarmyne et al., 2013). This starts with the difficulty in defining economic openness which is currently portrayed by the sum of all imports and exports (Zimmerman, 2017). The general idea about the growth and decline of corruption provides us with very little knowledge as well on how it affects the system of corruption. However, these relations are often taken as causations (as is done in the built simulation model). Still, there might be many more unknown variables involved. The more qualitative a system becomes or the more the system touches on the direct involvement of humans in that system the more uncertain the relations become and the higher the likelihood that the model will become invalid.

5.3.4. Modelling Choices

A specific modelling choice, made on real world assumed behaviour might by itself show correct behaviour. But a combination of several of these choices can cause static behaviour. For instance, the modelling choice on the relation between corruption and the level of free media. It is assumed in the model that when corruption rises free media will decline, based on a lookup function. However, as stated, these sub-systems are modelled as stocks with a single flow, making them very susceptible to exponential growth or decline. Whenever a certain level is reached it becomes more and more difficult to change the system (logarithmic) as assumed in the real world. This has been built into the sub-systems as well. Above a certain level of corruption another lookup function takes over that reduces the growth rate in order to balance the behaviour. This is done through the use of an if then else statement in the flow[1]. Both the use of if then else statements and lookups make a model very static and predictable. However, within the limited time frame of a masters thesis it is also impossible to establish relations between systems for which the relations are not known or of little significance.

The model behaviour shown in the level of democracy (autocracy democracy) for China and Russia (from stable behaviour to exponential growth to balancing behaviour) is expected to be initiated by an exponential growth of GDP. GDP is then affected by a growth factor, which is partly driven by the oil price. When the oil price rises, the GDP growth rate declines, when the oil prices fall the GDP growth rate rises (for oil exporting countries this is the other way around, however, it is assumed very likely that 'the people' will not profit from a rise in GDP in most OPEC countries). The oil price is based on historical data for half of the model run and further forecasted by a cross validation SD structure[2]. The model structure is shown in figure 5.5 and the GDP exponential growth (only for China and the United States in this specific model run) in figure 5.6.

[1]An if then else statement let you set a rule for a certain variable, if that rule is true, then x, if not true, than b. if p=y then x otherwise b
[2]Structure for cross validation on historical data has been retrieved from its modeller dr. (E) Erik Pruyt.

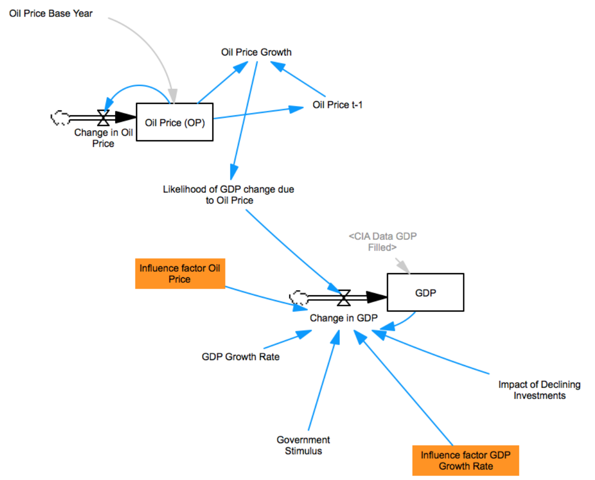


Figure 5.5: Influence of oil on GDP growth rate (model relations hidden for figure).

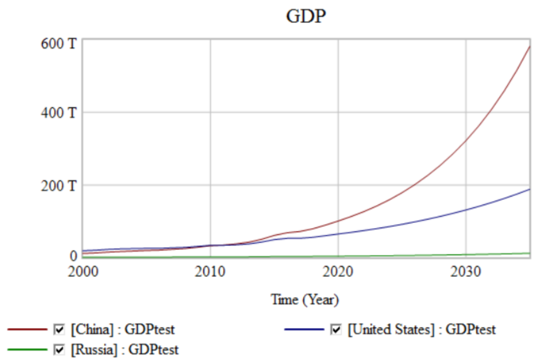


Figure 5.6: GDP outcome of a single model run

5.3.5. Missing or Lacking Data

As has been discussed already several times in this thesis, many flawed data sets are found. From large gaps between oil import and exports to tourist data that is only based on incoming and outgoing flights. However, datasets also often represent a particular situation. Relations between variables are often different in peace and war times. In a geopolitical model in which countries could fall in a state of conflict the change of the relations is, again, highly uncertain.

5.3.6. Conclusion

From the above it is concluded that the geopolitical simulation model built in this thesis is invalid and not useful for the understanding of geopolitical power shifts and therefore not useful in supporting scenario exploration with experts. This is, however, not driven by system dynamics as an approach, but mainly by the limited knowledge about the causal relations of the geopolitical system. In order to retrieve useful insights on a system, a good causal model is needed that calculates the relation between input and output variables. Besides that, the model should portray valid behaviour for new input values.

6

Conclusion and Discussion

6.1. Conclusion

This thesis explored and simulated the changes in power dynamics based on components of power by answering five questions: (1) What defines geopolitics? (2) What components does national power consist of? (3) Can these components be modelled? (4) How can these components be modelled? And, (5) is system dynamics a useful modelling approach in the modelling of geopolitics?

No other field has been more focused on shifts in world order than geopolitics. The ANV identifies geopolitics as the influence of geographical factors on issues of international politics. It is, however, found that geopolitics in the current world consists of more than just the influence of geographical factors. Therefore, the definition of geopolitics used in this thesis is broader and equals the broad term of international politics: Power, the most definite quantitative measurement of geopolitics. Three types of power have been identified: hard power (via coercion), soft power (via cooperation and attraction), and smart power (via strategies).

Hard power overwhelmingly takes the realist perspective and therefore its components reflect the same. Hard power: the power of coercion, relates to geographical factors, (natural) resources, population, economic power, and military strength.

Soft power, which refers to the co-optive use of power, rests on attractiveness and persuasion rather than coercion. Due to the positive notion to the definition of soft power it seems not to be clear where soft power starts and where it ends. This because, for instance, the line between persuasion and manipulation is very thin. For example, military power used for peacekeeping operations is seen as part of soft power although still coercive in nature or is part of hard power when taking the war theorist perspective of using good force against bad forces. Soft power structures were identified based on the Soft Power 30 index which was coined by Nye as the 'clearest picture of soft power to date' (McClory, 2017). In addition to the model-structures identified by using the index it was concluded that society should be part of a geopolitical model as well as the effect of society on the government in the pursuit of its goals can be significant. Unfortunately, it was also found that many of the used or described indices contain many measurement errors or that the methodologies are not published, making them difficult to understand and verify.

System dynamics, a modelling approach used for the modelling of complex non-linear systems is used to model the sub-systems identified as needed in the pursuit of simulating changes in power dynamics. Socio-economic systems, such as the geopolitical system are characterised by many deep uncertainties. Simulations of such complex systems are meant to support experts and should be used for exploratory purposes instead of forecasting. Models are useful as they increase our limited ability to perform scenario analysis.

It was found that modelling hard power is possible as the relations between the systems are often known, or known to a limited extent. Hard power model structures most often take real world quantities, making the modelling of hard power mainly difficult through lacking or missing data with which the system can not be initialised. However, it is found that the components of hard power cannot function properly by themselves and that they need softer components to complement them.

As soft power as a definition becomes more vague, its relations become more uncertain as well. This makes it more difficult to create quantitative structures for its measurement. Part of this uncertainty can be overcome by the use of exploratory modelling analysis (EMA). EMA provides the modeller and policy makers with clearer insights about the system and its behaviour. It is here where system dynamics and EMA complement each other. System dynamics, in which it is able to model complex systems and EMA which use can overcome the uncertainties within that system.

It was possible to build a geopolitical system dynamics model by using data for the initialisation of model structures, regression analysis to identify first order relations, and the modelling of empirical findings in the real world. However, the simulation model was found invalid as its behaviour over 500 simulation runs showed static behaviour assumed to be driven by explicit modelling choices. Choices that where logical, but in combination resulted in static model behaviour.

These modelling choices reflect a lack of knowledge about the relations between these complex systems, the nonexistence of needed data (for instance data on corruption), the misrepresentation of data due to measurement inconsistencies, and in the use of indices that in itself are already simplification of the system that often contain measurement inconsistencies, unpublished methodologies, and the use of indices as a metric for another index.

Finally, it has been concluded that system dynamics is an useful approach in the modelling of geopolitics and the simulation of national power shifts, but that potential future events are not possible to be simulated due to the complexity and uncertainties as described above. It is possible to simulate general characteristics of the system to provide a general idea about power shifts. If a good causal model, showing valid behaviour, is available for simulation and exploring scenario's. It will be an important tool in supporting experts in their scenario analysis.

In the following section a discussion is put forward on the future of modelling geopolitics and future perspectives are given.

6.2. Discussion

The conclusion of the research described in this is thesis is that modelling geopolitics with the goal of measuring national power is a very difficult exercise, but might show to be very useful. This discussion will focus on what is needed to take the next step in pursuing a good causal model.

The most urgent issue to be addressed is the limited knowledge on the causal relations between the systems. It is very unlikely that causal relations will be found using traditional methods due to the complexity. Future developments in artificial intelligence (AI) and machine learning could provide a better understanding on the causations between the variables within the system. However, machine learning asks for large amounts of data which is – at the moment of writing this thesis – not yet available.

However, large amounts of telecom data are available from which human decision making or behaviour can be derived. Shifting the modelling approach from relations between systems to the humans affecting the system, Bogomolov (2017), for instance, derives 70% accurate predictions from telecom data for crime intensive areas in metropoles in the next month. He also writes that social interactions could be traced and measured in relatively cheap way by the use of telecom data, using the combination of verbal and non-verbal behavioural patters. Bogomolov (2017) writes that by the use of 'affective states recognition' it is possible to derive a data-driven model of individual human behaviour from telecom data. The examples provided for the use of machine learning are extremely specific. The interactions and influences of humans on the shaping of the geopolitical system might be worthwhile researching as for instance private investments (important for the growth or decline of GDP) are driven by a qualitative believe in the stability of an economy. Further research should therefore be conducted in machine learning for the use of establishing causal relations for the modelling of geopolitics.

Limitations of geopolitical models will still lie in the ability of understanding the system and the making of generalisations in the relations. More data will become available, and AI will push the accuracy of those generalisations. A few limitations within the modelling itself are present as many different modelling techniques exist that could complement each other in order to converge to a representation of the real system.

6.3. Contributions of the Thesis

This thesis contributed to both the research fields of international relations as to the modelling in that it is – by the knowledge of the author – the first geopolitical model built for the analysis of state power dynamics to perform uncertainty analysis using EMA and the EMA-Workbench. It is also the first model that took into account cultural databases for the modelling of societal behaviour. It is also the first research reflecting extensively on the literature of national power and complements this with a dynamic simulation model.

6.4. Future Perspectives

Further research should mainly be conducted in finding causal relations that are useful in the general modelling of geopolitics. There is no need for very specific relations as the whole system is too large to be modelled on a very low level of resolution. As discussed above machine learning can be of help.

Moreover, more research is needed in complementing various modelling approaches in the pursuit of simulating geopolitics. Although possible in system dynamics, Agent Based Modelling (ABM) instead of the system, single entities and their behaviour are modelled. This might be useful in the inclusion of global engagement as ABM provides more possibilities in the establishments or break-ups of coalitions.

Another field for further research is the inclusion of Non-Governmental Organisations (NGO's). NGO's gain more power and are more likely to influence the international political arena (Nye, 2004).

Finally, further research should be focused on a more exploratory analysis on geopolitical models and should interpret the results through the EMA-Workbench in identifying which relations are dominant in the model.

References

Adhikari, S. (2013). Modern geopolitics versus Post-Modern geopolitics: a critical review. *Transactions*, *35*(1), 35–48.

Al-Rodhan, N. (2014). *Meta-Geopolitics: the Relevance of Geopolitics in the Digital Age.* Retrieved 2018-06-15, from https://www.e-ir.info/2014/05/25/meta-geopolitics-the-relevance-of-geopolitics-in-the-digital-age/

Analistennetwerk Nationale Veiligheid. (2016). *Nationaal Veiligheidsprofiel 2016* (Tech. Rep.). Bilthoven: RIVM.

Bankes, S. (1993). Exploratory Modelling for Policy Analysis. *Operations Research*, *41*(3), 435–449. doi: 10.1287/opre.41.3.435

Bavardage Diplomatique. (2016, aug). *Hard Power and Diplomacy.* Retrieved 2018-06-19, from https://bavardagediplomatiqueblog.wordpress.com/2016/08/30/42/

BC Patient Safety and Quality Council. (2018). *About the Council.* Retrieved from https://bcpsqc.ca/about-the-council/

Bennet, D. (2017). *What Causes GDP Decline?* Retrieved August 15, 2018, from https://bizfluent.com/info-8611383-causes-gdp-decline.html

Bogomolov, A. (2017). *Predictive Modeling of Human Behavior: Supervised Learning from Telecom Metadata* (PhD). Trento University.

Brookings Institution. (2017). *Korean unification: Prospect and global implications.* Retrieved from https://www.brookings.edu/events/korean-unification-prospect-and-global-implications/

Bull, H. (1966). International Theory: The Case for a Classical Approach. *World Politics*, *18*(3), 361–377.

Carr, E. H. (1946). *The Twenty Years' Crisis 1919-1939* (2nd ed.). London: MacMillan.

Chan, K. (2016, dec). *These are the most powerful languages in the world.* Retrieved 2018-07-03, from https://www.weforum.org/agenda/2016/12/these-are-the-most-powerful-languages-in-the-world/

Chang, C.-L. (2004). A Measure of National Power. *International seminar at the National University of Malaysia*, 34. Retrieved from http://www.fgu.edu.tw/{~}academic/up1/malayconference1/paperscollection.files/conferefncepapers/chang.pdf

Chang, C.-L. (2014). *A Measure of National Power* (Tech. Rep.). Fo-guang University.

Cholley, A. (1942). *Guide de l'Etudiant en Geographie.* Paris: Presses Universitaires de France.

Cline, R. S. (1975). *World Power Assessment: A Calculus of Strategic Drift.* Center for Strategic and International Studies, Georgetown University.

Cohen, S. (1964). *Geography and Politics in a Divided world.* London: Methuen & Co.

Dugin, A. (1997). Foundations of Geopolitics: Russia's Geopolitical Future.

Efird, B., & Genna, G. M. (2002). Structural Conditions and the Propensity for Regional Integration. *European Union Politics*, *3*(3), 267–295. doi: 10.1177/1465116502003003001

Ernst \& Young. (2012). *Rapid-growth markets soft power index* (Tech. Rep.). Ernst \& Young.

European Parliament. (1997). The Role of the European Union and Sport. In *Power shifts in a changing world order.*

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review*, *97*(01), 75–90. Retrieved from http://www.journals.cambridge.org/abstract{_}S0003055403000534 doi: 10.1017/S0003055403000534

Forrester, J. W. (1968). *Principles of Systems* (Vol. 2) (No. Chapter 6). Cambridge: Wright-Allen Press. Retrieved from http://www.amazon.com/Principles-Systems-Jay-Wright-Forrester/dp/1883823412 doi: 10.1007/978-1-4939-1920-8

Fraser, M. (2009). *Geopolitics 2.0.* doi: http://www.realinstitutoelcano.org/wps/portal/rielcano_en/contenido?WCM_GLOBAL_CONTEXT=/elcano/elcano_in/zonas_in/defense+security/ari144-2009

Freedom House. (2018). *Freedom in the World 2018: Methodology.* Retrieved 2018-07-03, from https://freedomhouse.org/report/methodology-freedom-world-2018

Global Fire Power. (2018). *Global Fire Power Index.* Retrieved 2018-06-30, from https://www.globalfirepower.com/

Gould, R. V. (2002). The Origins of Status Hierarchies: A Formal Theory and Empirical Test. *American Journal of Sociology*, *107*(5), 1143–1178. Retrieved from http://www.journals.uchicago.edu/doi/10.1086/341744 doi: 10.1086/341744

GR:EEN. (2014, jun). *Project context and objectives.* Author. Retrieved from https://warwick.ac.uk/fac/soc/pais/research/researchcentres/csgr/green/work{_}packages/context/

Halevy, N., Y. Chou, E., & D. Galinsky, A. (2011, feb). A functional model of hierarchy. *Organizational Psychology Review*, *1*(1), 32–52. Retrieved from http://journals.sagepub.com/doi/10.1177/2041386610380991 doi: 10.1177/2041386610380991

Hart, J. (1976). Three approaches to the measurement of power in international relations. *International Organization.*

Herwig, H. H. (1999). Geopolitik: Haushofer, Hitler and Lebensraum. *Journal of Strategic Studies*, *22*(2-3), 218–241. Retrieved from http://shodhganga.inflibnet.ac.in/handle/10603/18345 doi: 10.1080/01402399908437762

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and Organizations: Software of the Mind.* McGraw-Hill.

Höhn, K. (2011). *Geopolitics and the Measurement of National Power* (Unpublished doctoral dissertation). Hamburg University.

Hughes, B. B. (2009). *Forecasting Long-Term Global Change: Introduction to International Futures (IFs)* (No. December).

Hughes, B. B. (2014). *IFs Population Model Documentation.*

Huh, M. Y., Cha, M. S., Jeong, Y. T., Jung, H. S., Kang, G. S., Kwon, O. K., ... Yeo, I. K. (2012). *Basic Reading on Korean Unification Basic Reading on Korean Unification* (Tech. Rep.). Seoul: Korea Institute for National Unification.

Ilgen, T. L. (2006). *Hard power, soft power, and the future of transatlantic relations.* Ashgate Publishing Limited.

In.Europa. (2017a). *How we describe state power.* Retrieved 2018-06-15, from http://index.ineuropa.pl/en/measures-of-state-power/how-we-describe-state-power/

In.Europa. (2017b). *Key conclusions.* Retrieved 2018-06-15, from http://index.ineuropa.pl/en/key-conclusions/

Johnston, J. (1963). *Econometric Methods* (McGraw-Hill, Ed.). New-York.

Jones, S. E. (2006). *Against Technology: From the Luddites to Neo-Luddism.* Taylor & Francis.

Kasperson, R. E., & Minghi, J. V. (1969). *The structure of political geography.*

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2010). *The Worldwide Governance Indicators: Methodology and Analytical Issues* (Tech. Rep.). The World Bank.

Keen, D. (2008a). Combatants and Their Grievances. In *Complex emergencies* (pp. 50–71). Cambridge: Polity.

Keen, D. (2008b). Defining the Enemy. In *Complex emergencies* (pp. 71–99). Cambridge: Polity.

Keen, D. (2008c). 'Greed': Economic Agenda. In *Complex emergencies* (pp. 25–50). Cambridge: Polity.

Kekic, L. (2007). *The Economist Intelligence Unit's index of democracy* (Tech. Rep.).

Kendall-Taylor, A., & Frantz, E. (2016). *Why Populism Is a Pathway to Autocracy.* Retrieved 2018-05-21, from https://www.foreignaffairs.com/articles/2016-12-05/how-democracies-fall-apart

Kennedy, P. (1987). *The Rise and Fall of the Great Powers.* New York: Random House.

Kim, J.-O., Kim, P., Hassanein, A. M., Kim, S., Lee, M.-j., Wang, J., ... Kim, M. (2012). *Index of National Power.* Retrieved 2018-06-30, from http://www.nationalpower.info/

Kjéllen, R. (1916). *Staten som lifsform.* Stockholm: Hugo Grebers forlag.

Koski, J. E., Xie, H., & Olson, I. R. (2015). Understanding social hierarchies: The neural and psychological foundations of status perception. *Social Neuroscience*, *10*(5), 527–550. doi: 10.1080/17470919.2015.1013223

Kugler, J., & Organski, K. (1989). The Power Transition: A Retrospective and Prospective Evaluation. *Handbook of War Studies*, 171–194.

Kwakkel, J. H. (2017). *EMA Workbench.* Retrieved August 16, 2018, from https://github.com/quaquel/EMAworkbench

Lackey, D. P. (2015). Soft Power, Hard Power, And Smart Power. *The Philosophical Forum, Inc.*.

Lange, G.-M., Wodon, Q., & Carey, K. (2018). *The Changing Wealth of Nations* (Tech. Rep.). Washington: World Bank Group.

Lee, M. (2017). The EU and the Six-Party Talks. In N. Casarini (Ed.), *Promoting security cooperation and trust building in northeast asia: The role of the european union* (1st ed., pp. 169–182). Rome: Edizioni Nuova

Cultura.
Lempert, R. J., Popper, S., & Bankes, S. (2003). *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis* (Tech. Rep.).
Levin, K., Cashore, B., Bernstein, S., & Auld, G. (2012). Overcoming the tragedy of super wicked problems: Constraining our future selves to ameliorate global climate change. *Policy Sciences*, *45*(2), 123–152. doi: 10.1007/s11077-012-9151-0
Lutmer-Paulson, M. (2017). *What are the Causes of Economic Decline.* Retrieved August 15, 2018, from https://classroom.synonym.com/what-are-the-causes-of-economic-decline-12084036.html
Mamadouh, V. D. (1998). Geopolitics in the nineties: One flag, many meanings. *GeoJournal*, *46*(4), 237–253. doi: 10.1023/A:1006950931650
Martin, W. (2017a, sep). *This chart shows what a war with North Korea could do to the global economy.* Retrieved from https://www.businessinsider.nl/economic-impacts-of-war-between-usa-north-korea-2017-9/?international=true{&}r=UK
Martin, W. (2017b, aug). *What a war between North Korea, US could do to global economy - Business Insider.* Retrieved from http://uk.businessinsider.com/war-between-north-korea-and-usa-global-economy-impact-2017-8
Mazur, A. (2013). Biosocial Model of Status in Face-to-Face Primate Groups. *Procedia - Social and Behavioral Sciences*, *84*(2), 53–56. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S1877042813015759 doi: 10.1016/j.sbspro.2013.06.508
McClory, J. (2017). *The Soft Power 30: A Global Ranking of Soft Power* (Tech. Rep.). Portland: USC Center on Public Deplomacy.
McClory, J. (202). *The New Persuaders III: A 2012 Global Ranking of Soft Power* (Tech. Rep.). Institute for Government.
Meadows, D., Richardson, J., & Bruckmann, G. (1982). *Groping in the Dark: The First Decade of Global Modelling* (Vol. 33) (No. 12). Bristol: International Institute for Applied System Analysis. doi: 10.1057/jors.1982.234
Millennium Institute. (2018). *T21 Model Structure.* Retrieved from http://changingcourse-agriculture.com/workstream/t21-model-planning/t21-model-structure/
Mogherini, F. (2017). *"Europe and Asia - building a Cooperative Global Order". Speech by the HR/VP Federica Mogherini at Tsinghua University* (Tech. Rep.). Brussels: European Union External Action Service.
Morgenthau, H. J. (2001). *Politics Among Nations. The Struggle for Power and Peace* (sixth edit ed.; K. W. Thompson, Ed.). New Delhi: Kalyani. Retrieved from http://www3.nd.edu/{~}cpence/eewt/Morgenthau2005.pdf doi: 10.2307/2086875
Mullan, T. L. B. M. (2016, jul). *International (Non-)Cooperation in a Changing World Order.* Council on Foreign Relations. Retrieved from https://www.cfr.org/blog/international-non-cooperation-changing-world-order
Nersessian, N. J., & MacLeod, M. (2017). Models and Simulation. In L. Magnani & T. Bertolotti (Eds.), *Handbook of model-based science* (p. 1179). Pavia: Springer.
Norris, R., Norris, B., & Haring, L. (1980). *Political Geography.* Columbus: Charles E Merrill Publishing Co.
Nye, J. (2004). Soft power: The means to success in world politics. In *Fan, y., soft power: Power of attraction or confusion* (Vol. 2, pp. 99–100). Retrieved from http://books.google.com/books?hl=en{&}lr={&}id=sFNfYvNtw5AC{&}oi=fnd{&}pg=PR9{&}dq=Soft+Power:+The+Means+to+Success+in+World+Politics{&}ots=7q{_}Wgd47Pj{&}sig=Who4NAkLyEW1AuvbonhtCnh7ERk doi: 10.2307/1148580
Nye, Jr., J. S. (1990). Soft Power. *Foreign Policy*, *80*(80), 153–171. doi: 10.1017/S1052703600007164
Nye, Jr., J. S. (2003). *Soft power : Propaganda isn't the way.* Retrieved 2018-06-30, from https://www.nytimes.com/2003/01/10/opinion/soft-power-propaganda-isnt-the-way.html
Nye, Jr., J. S. (2006, feb). *Think Again: Soft Power.* Retrieved 2018-07-03, from https://foreignpolicy.com/2006/02/23/think-again-soft-power/
Nye, Jr., J. S. (2009, aug). *Get Smart: Combining hard and Soft Power.* Retrieved 2018-07-03, from https://www.foreignaffairs.com/articles/2009-07-01/get-smart
Nye, Jr., J. S. (2011). *The Future of Power* (First ed.). New York: Perseus Books Group.
OECD. (2018). *Trust in Government.* Retrieved 2018-07-03, from http://www.oecd.org/gov/trust-in-government.htm
Organski, A. (1958). *World Politics.* New York: Knopf.

O'Tuathail, G. (2005). *Critical Geopolitics: The Politics of Writing Global Space.* London: Taylor & Francis e-Library. doi: 10.1017/CBO9781107415324.004

Pilling, D. (2018a). *The Growth Delusion.* New York: Tim Duggan Books.

Pilling, D. (2018b). *Why it is time to change the way we measure the wealth of nations.* Retrieved 2018-06-15, from https://www.ft.com/content/765781fe-f136-11e7-b220-857e26d1aca4

Pollack, J. D. (2001). Korean Unification: Illusion or Aspiration? *The Brown Journal of World Affairs*, *8*(1), 77–90.

Preston, S. H. (1975). The Changing Relation between Mortality and level of Economic Development. *Population Studies*, *29*(2), 231–248. doi: 10.1080/00324728.1975.10410201

Pruyt, E. (2013). *Small System Dynamics Models for Big Issues: Triple Jump towards Real-World Complexity.* Delft: TU Delft Library.

Pruyt, E. (2014). *From Data-Poor to Data-Rich.*

Quarmyne, R., William, C., & Arhenful, P. (2013). Determining Factors Affecting Corruption : A Cross Country Analysis. , *1*(10), 275–285.

Ratzel, F. (1923). *Politische Geographie.* M\"unchen.

Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155–169. Retrieved from http://www.springerlink.com/content/m5050140x48140m3/{%}5Cnhttp://dx.doi.org/10.1007/BF01405730 doi: 10.1007/BF01405730

Roman, P., & Thiry, G. (2016). The inclusive wealth index. A critical appraisal. *Ecological Economics*, *124*, 185–192. doi: 10.1016/j.ecolecon.2015.12.008

Rummel, R. (1976). *Understanding Conflict and War. Volume 2: The Conflict Helix.* New York: John Wiley & Sons.

Sahoo, D. K. (2010). *A Geopolitical Perspective on Central Asia-China Relations* (Tech. Rep.). Jawahrlal Nehru University.

Saradzhyan, S., & Abdullaev, N. (2018). *Is Vladimir Putin's Russia in decline? We figured out how to measure 'national power.'.* Retrieved August 16, 2018, from https://www.washingtonpost.com/news/monkey-cage/wp/2018/06/20/is-vladimir-putins-russia-in-decline-we-figured-out-how-to-measure-national-power/?utm{_}term=.ab10b21397c2

Savin-Williams, R. C. (1979, dec). Dominance Hierarchies in Groups of Early Adolescents. *Child Development*, *50*(4), 923. Retrieved from http://www.jstor.org/stable/1129316?origin=crossref doi: 10.2307/1129316

Schwab, K. (2018). *The Global Competitiveness Report 2017-2018* (Tech. Rep.). World Economic Forum.

Shell Global. (2016). *World Energy Model: A View to 2100* (Tech. Rep.).

Shell Global. (2018). *What Are Shell Scenarios?* Retrieved 2018-06-19, from https://www.shell.com/energy-and-innovation/the-energy-future/scenarios/what-are-scenarios.html

Silbereisen, R. K. (2016). Psychological challenges of unification - Selected results and thoughts on Korea. *Historical Social Research*, *41*(3), 299–317. doi: 10.12759/hsr.41.2016.3.299-317

Smitha, N. (2013, aug). *Facebook Metrics Defined: Engagement Rate.* Retrieved 2018-07-03, from https://simplymeasured.com/blog/facebook-metrics-defined-engagement-rate/{#}sm.000cwmk7ubjheqe10uc2plm1do49f

Spykman, N. J. (1938). Geography and Foreign Policy, II. *The American Political Science Review*, *32*(2), 213–236. Retrieved from http://www.jstor.org/stable/1948667{%}5Cnhttp://about.jstor.org/terms doi: 10.2307/1948667

Spykman, N. J. (1942). *America's Strategy in World Politics, the United States and the Balance of Power.* Transaction Publishers.

Steinmetz, S. (1929). *Soziologie des Krieges* (J. A. Barth, Ed.). Leipzig.

Sterman, J. D. (2000). *System Thinking and Modeling for a Complex World.* McGraw-Hill.

The Economist. (2012, jun). *The real wealth of nations.* Retrieved 2018-06-22, from https://www.economist.com/node/21557732

The Economist. (2014a, nov). *How to lie with indices.* Retrieved 2018-07-07, from https://www.economist.com/international/2014/11/08/ranking-the-rankings,https://www.economist.com/leaders/2014/11/08/how-to-lie-with-indices

The Economist. (2014b, nov). *Ranking the rankings.* Retrieved 2018-07-03, from https://www.economist.com/international/2014/11/08/ranking-the-rankings

The Fund For Peace. (2018). *2018 Fragile State Index* (Tech. Rep.). The Fund For Peace.

The Heritage Foundation. (2018). *Index of Economic Freedom: Methodology* (Tech. Rep.).

Transparency International. (2017). *Corruption Perception Index.* Retrieved August 16 2018, from https://www.transparency.org/news/feature/corruption{_}perceptions{_}index{_}2017

Turner, S., Neill, C., Hughes, B. B., & Narayan, K. (2017). *Guide to scenario analysis in international futures (IFs).* Denver.

UNEP, & UNU-IHDP. (2012). *Inclusive Wealth Report 2012: Measuring progress toward sustainability* (Tech. Rep.).

UNEP, & UNU-IHDP. (2014). *Inclusive Wealth Report 2014: Measuring progress toward sustainability* (Tech. Rep.). UNEP.

United Nations. (2018a). *E-Government Development Index (EGDI).* Retrieved 2018-07-03, from https://publicadministration.un.org/egovkb/en-us/About/Overview/-E-Government

United Nations. (2018b). *E-Participation Index.* Retrieved 2018-07-03, from https://publicadministration.un.org/egovkb/en-us/About/Overview/E-Participation

United Nations. (2018c). *Gender Inequality Index (GII).* Retrieved 2018-07-03, from http://hdr.undp.org/en/content/gender-inequality-index-gii

United Nations. (2018d). *Human Development Index (HDI).* Retrieved 2018-07-03, from http://hdr.undp.org/en/content/human-development-index-hdi

Ventana Systems Inc. (2011). *Vensim Systems Manual* (Tech. Rep.).

Vinci, A. (2006). Greed-Grievance Reconsidered: The Role of Power and Survival in the Motivation of Armed Groups. *Civil Wars*, *8*(1), 25–45. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/13698240600886032 doi: 10.1080/13698240600886032

Vogel, L., & Best, H. (2016). Political Elites in Transition and Unification: German Lessons for the Korean Peninsula? *Historical Social Research*, *41*(3), 336–367.

Vogt, M., Bormann, N. C., Rüegger, S., Cederman, L. E., Hunziker, P., & Girardin, L. (2015). Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family. *Journal of Conflict Resolution*, *59*(7), 1327–1342. Retrieved from http://www.epr.ucla.edu/ doi: 10.1177/0022002715591215

Walker, W. E., Lempert, R. J., & Kwakkel, J. H. (2012). *Deep Uncertainty.*

Wigman, S. (n.d.). *Global Migration Dynamics: An Exploratory Study Integrating Multi-Resolution Modeling Techniques with Semiautomatic Data Acquisition* (Masters). Technical University Delft.

without borders, R. (2018). *Detailed Methodology.* Retrieved 2018-07-03, from https://rsf.org/en/detailed-methodology

World Justice Project. (2018). *Rule of Law Index.* Retrieved August 16 2018, from http://data.worldjusticeproject.org

Yan, X. (2006). The rise of China and its power status. *Chinese Journal of International Politics*, *1*(1), 5–33. doi: 10.1093/cjip/pol002

Yesilada, B. A., Kugler, J., Genna, G. M., & Tanrikulu, O. G. (2017). *Global power transition and the future of the European Union.* Retrieved from https://www.amazon.com/Transition-European-Routledge-Contemporary-Studies/dp/1138283495/ref=sr{_}1{_}1?s=books{&}ie=UTF8{&}qid=1513719958{&}sr=1-1{&}keywords=Global+power+transition doi: 10.4324/9781315270227

Young C. Kim. (2013). the United States and Korean Reunification. In J. H. Bae (Ed.), *Korea observer* (Vol. 10, pp. 115–121). Seoul: Korean Institute for National Unification.

Zakaria, F. (2017). *The decline of U.S. influence is the great global story of our age.* Retrieved August 16, 2018, from https://www.washingtonpost.com/opinions/global-opinions/the-decline-of-us-influence-is-the-great-global-story-of-our-times/2017/12/28/bfe48262-ebf6-11e7-9f92-10a2203f6c8d{_}story.html?noredirect=on{&}utm{_}term=.2081b7555622

Zimmerman, C. (2017). *Are You Open?* Retrieved August 16 2018, from https://fredblog.stlouisfed.org/2017/05/are-you-open/

Appendices

A

Tables Components Soft Power

A.1. Culture

Table A.1: Components of sub-index 'Culture' with variable type and comments from Soft Power 30 index (McClory, 2017)

Metric	Data Source	Type	Comment
Culture			
Total number of tourist arrivals	UN World Tourism Organisation	Stock	-
Average spend per tourist (total tourism receipts divided by number of tourists)	UN World Tourism Organisation	Auxilary variable	-
Number of films appearing in major film festivals	Various	Stock	-
Number of foreign correspondents in the country	Gorkana Media Database / Foreign Correspondent Associations / Various	Stock	-
Number of UNESCO World Heritage sites	UNESCO Statistics	Stock	-
Annual museum attendance of global top 100	The Art Newspaper Review Number 289, April 2017	Flow	-
Size of music market	IFPI Global Music Report 2017 IFPI Global Music Report 2017	Index	-
Number of top 10 albums in foreign countries	IFPI Global Music Report 2017 IFPI Global Music Report 2017	Stock	-
Number of top 10 albums in foreign countries	IFPI Recording Industry in Numbers 2016	Stock	-
Olympic medals (Summer 2016 / Winter 2014)	International Olympic Committee	Stock	-
FIFA Ranking (Men?s)	FIFA/Coca Cola World Rankings	Index	-
Quality of national air carrier	Skytrax Arline Quality Review	Index	-
Michelin starred restaurants	Michelin guide	Index	-
Power Language Index (PLI)	Chan, K., Power Language Index, 2016	Index	Measures how powerful a language is based on: geography (countries spoken), economy (GDP), communication (native speakers), knowledge/media (internet content, academic journals), and diplomacy (official languages from institutions such as UN) (Chan, 2016).

A.2. Global Engagement

Table A.2: Components of sub-index 'Global Engagement' with variable type and comments from Soft Power 30 index (McClory, 2017)

Metric	Data Source	Type	Comment
Global Engagement			
Total overseas development aid	OECD	Auxiliary	-
Overseas development aid / GNI	OECD	Stock	-
Number of embassies abroad	Lowy Institute / Embassypages / Various	Stock	-
Number of embassies in the country	Embassypages	Stock	-
Number of consulates general abroad	Lowy Institute / Embassypages / Various	Stock	-
Number of permanent missions to multilateral organisations	Lowy Institute / Various	Stock	-
Membership of international organisations	Various	Stock	-
Environmental treaty signatures	United Nations Treaty Collection	Stock	-
Asylum seekers per 1,000 people	World Bank / Asylum Seeker Resource Centre	Stock	-
Number of diplomatic cultural missions	Various	Stock	-
Number of countries a citizen can visit visa-free	Henley & Parners Visa Restrictions Index 2016	Stock	-
Size of Weekly Audience of State Broadcaster	Various	Stock	-
Environmental Performance Index (EPI)	Yale Center of Environmental Law & Policy (YCELP)	Index	-

A.3. Government

Table A.3: Components of sub-index 'Government' with variable type and comments from Soft Power 30 index (McClory, 2017): Part 1

Metric	**Data Source**	**Type**	**Comment**
Government			
Human Development Index score	UNDP Human Development Repo	Index	The Human Development Index (HDI) was established to accentuate that the capabilities of people is the final benchmark for assessing a countries development (United Nations, 2018d). Life expectancy, expected years of schooling, mean years of schooling, and GDP per capita are taken to calculate the ranking of the HDI (United Nations, 2018d).
Freedom House Index score	Freedom House	Index	The Freedom House Index reflects on a countries political and civil freedom. Political freedom takes 40% of the total score and consists of: electoral process, political pluralism and participation, and functioning of government. Whereas civil freedom represents 60% containing: freedom of expression and beliefs, associational and organisational rights, rule of law, and personal autonomy and individual rights (Freedom House, 2018).
Number of think tanks in the country	McGann, J. (2017), 2016 Global Go to Think Tank Index Report	Stock	Not been taken into account in the stock-flow structure as, to me, it is unclear on which basis new think tanks arise except for a relative position to other countries number or a need for information and thus subsidisation from the government.
Gender Equality Index score	UNDP Human Development Report	Index	The Gender Inequality Index (GII) bases its score on the following indicators: maternal mortality ratio, adolescent birth rate, female and male population with at least secondary education, female and males share of parliamentary seats, and female and male labour force participation rates (United Nations, 2018c).
Economist Democracy Index score	Economist Intelligence Unit	Index	In the Democracy Index the following indicators have been taken into account: electoral process and pluralism, functioning of government, political participation, democratic political structure, and civil liberties (Kekic, 2007)

Table A.4: Components of sub-index 'Government' with variable type and comments from Soft Power 30 index (McClory, 2017) :Part 2

Metric	Data Source	Type	Comment
Government			
Size of shadow economy as a percentage of GDP[1]	Hassan, M & Schneider, F (2016), Size and Development of the Shadow Economies of 157 Countries Worldwide: Updated and New Measures from 1999 to 2013	Flow	Has not been taken into account in the stock-flow structure as I do not see this as relevant in the perception of power and in the wide inclusion of other components.
Homicides per capita	World Bank	Auxiliary	-
World Bank Voice and Accountability Index score	World Bank	Index	"capturing perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media" (Kaufmann, Kraay, & Mastruzzi, 2010)
Capital punishment carried out in 2016	Amnesty International	Stock	-
Income inequality - gini coefficient	World Bank	Auxiliary	-
World Economic Forum Trust in Government Index score	World Economic Forum	Index	The trust in a government has been measured over: reliability (minimise uncertainty in economic, social and political environment), responsiveness (public services), openness (access to information), better regulation (justice, rule of law), Integrity & fairness (corruption), and inclusive policy making (OECD, 2018).
Press Freedom Index score	Reporters Without Borders	Index	The Press Freedom Index captures press freedom in: pluralism, media independence, environment and self-censorship, legislative framework, transparency, infrastructure, and abuses (without borders, 2018).
World Bank Government Effectiveness score	World Bank	Index	"Capturing perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies" (Kaufmann et al., 2010).
World Bank Good Governance Regulation Quality score	World Bank	Index	"capturing perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development" (Kaufmann et al., 2010).
World Bank Good Governance Rule of Law score	World Bank	Index	"capturing perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence" (Kaufmann et al., 2010).

A.4. Education

Table A.5: Components of sub-index 'Education' with variable type and comments from Soft Power 30 index (McClory, 2017)

Metric	Data Source	Type	Comment
Education			
Average of OECD PISA science, maths and reading scores	OECD	Stock	-
Gross tertiary educational enrolment rate	World Bank	Rate	-
Number of top global universities	Times Higher Education (top 200)	Stock	-
Number of academic science journal articles published	World Bank	Stock	Can be easily boosted by publishing rubbish articles. Not a robust measurement.
Number of international students in the country	UNESCO Institute for Statistics	stock	-
Spending on education as percentage of GDP	World Bank	Stock	-

A.5. Digital

Table A.6: Components of sub-index 'Digital' with variable type and comments from Soft Power 30 index (McClory, 2017)

Metric	Data Source	Type	Comment
Digital			
Facebook followers for heads of state (outside of country)	Facebook	Stock	-
Facebook engagement score for heads of state or government (outside of country)	Auxiliary	-	
Facebook followers for ministry of foreign affairs (outside of country)	Facebook	Stock	-
Facebook engagement score for ministry of foreign affairs (outside of country)	Facebook	Auxiliary	-
Number of internet users per 100 inhabitants	World Bank	Stock	-
Secure internet servers per 1 million people	World Bank	Stock	-
Mobile phones per 100 people	International Telecommunication Union	Stock	-
Internet bandwidth thousands Mpbs	International Telecommunication Union	Stock	-
Government Online Services Index	United Nations E-Government Survey	Index	Measures: provision of online services, telecommunication connectivity, and human capacity (United Nations, 2018a). It is extremely unclear what this index exactly does. Therefore, it has not been taken into account.
E-participation Index	Web Index	Index	This should be in the sub-index government as it measures the participation and openness of a government, but on a digital level (United Nations, 2018b).
Fixed broadband subscriptions per 100 people	World Bank	Stock	-

A.6. Enterprises

Table A.7: Components of sub-index 'Enterprises' with variable type and comments from Soft Power 30 index (McClory, 2017)

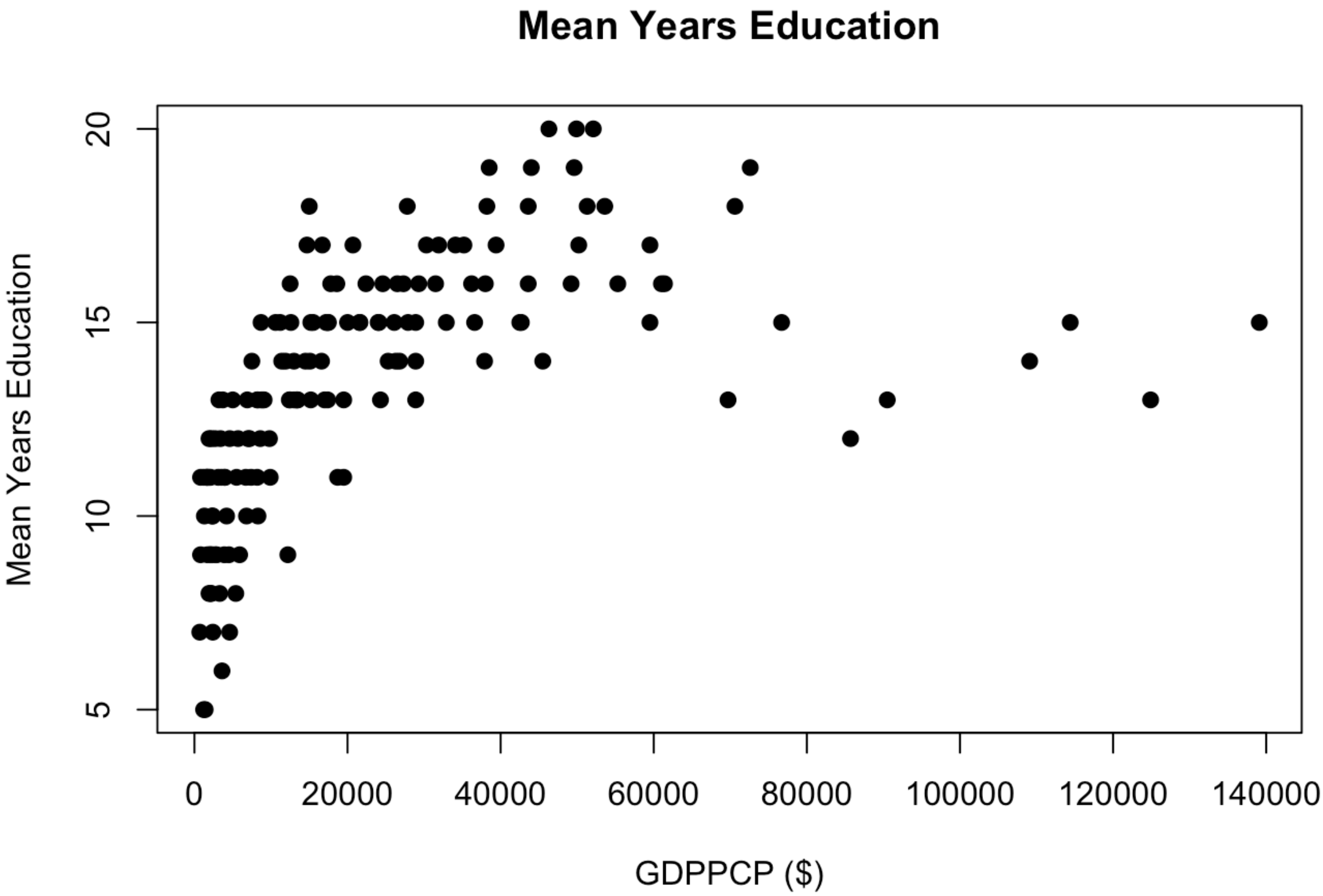
Metric	Data Source	Type	Comment
Enterprise			
Global patents filed (percentage of GDP)	World Intellectual Property Organization / World Bank	Flow	Depreciation of patents has to be taken into account. Therefore, this measurement is not correct. Besides that, the valuing of patents is very value based and can be mis-valued. Finally, this component can easily be boosted by countries by the filing of useless patents.
WEF Competitiveness Index	World Economic Forum	Index	competitiveness is seen as the set of institutions, policies, and factors that determine the level of productivity of an economy, which in turn sets the level of prosperity that the economy can achieve (Schwab, 2018)
Foreign direct investment as percentage of GDP	United Nations Conference on Trade and Development Statistics / World Bank / Various	Flow	-
Heritage Economic Freedom Index score	2017 Index of Economic Freedom	Index	In the Economic Freedom Index four groups of categories have been taken into account: rule of law (property rights, judicial effectiveness, and government integrity), government size (tax burden, government spending, fiscal health), regulatory efficiency (business freedom, labour freedom, and monetary freedom), and market openness (trade freedom, investment freedom, and financial freedom) (The Heritage Foundation, 2018)
Corruption Perceptions Index score	Transparency International Corruption Perceptions Index 2016	Index	-
R&D spending as a percentage of GDP	World Bank	Flow	-
Global Innovation Index score	The Global Innovation Index 2016	Index	-
Number of SMEs as a percentage of labour force working in SMEs	International Finance Corporation	Flow	-
World Bank Ease of Doing Business Report	World Bank	Index	-
Unemployment rate as a percentage of labour force	World Bank	Rate	-
Hi-tech exports as a percentage of manufactured exports	World Bank	Flow	-
Log of business start-up costs as a percentage of GNI per capita	World Bank	Flow	-

B
Regression Scripts

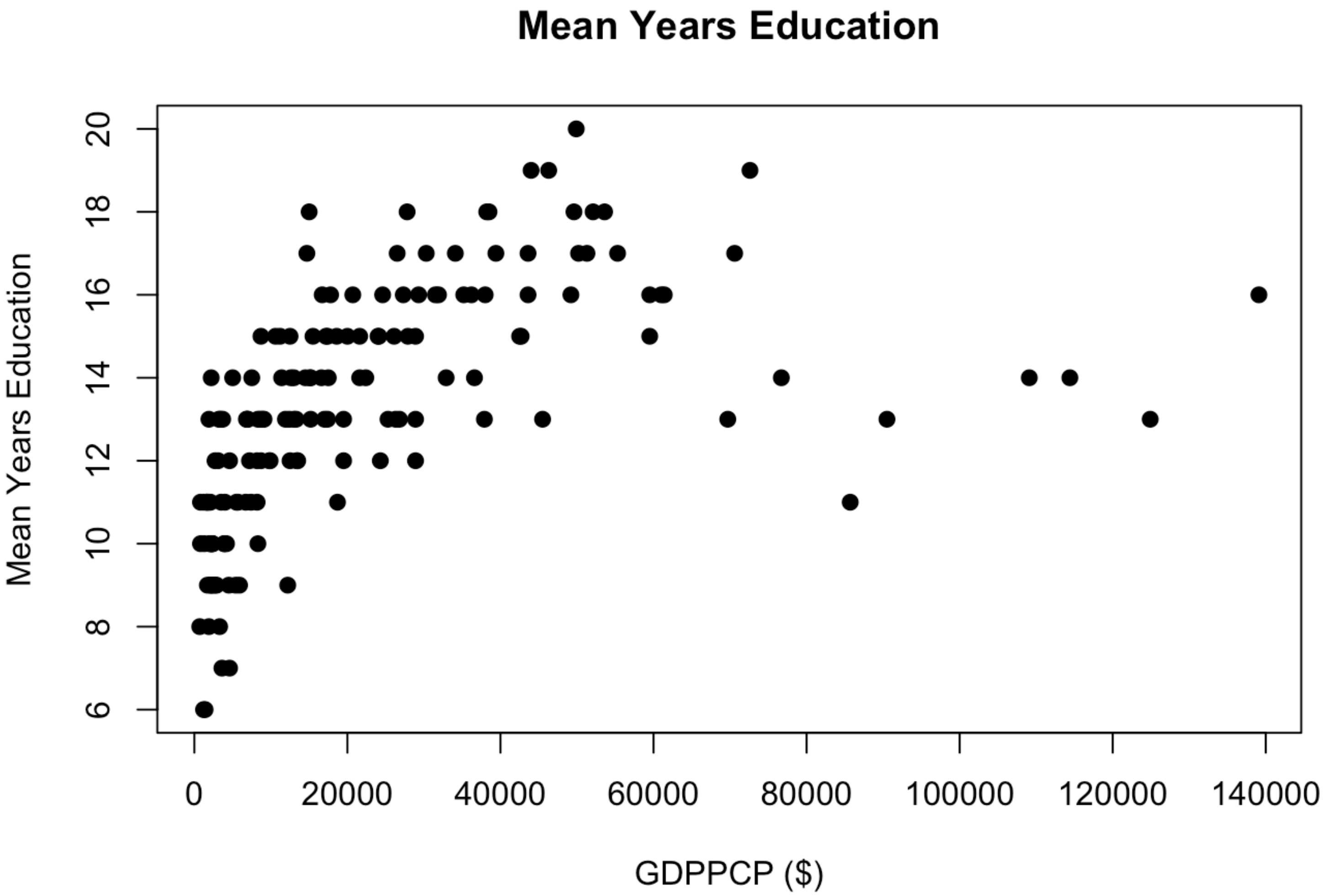
B.1. Mean Years of Education

Notebook Thesis Rémon ten Bhömer TU Delft 2018: Regression Curve Mean Years of Education

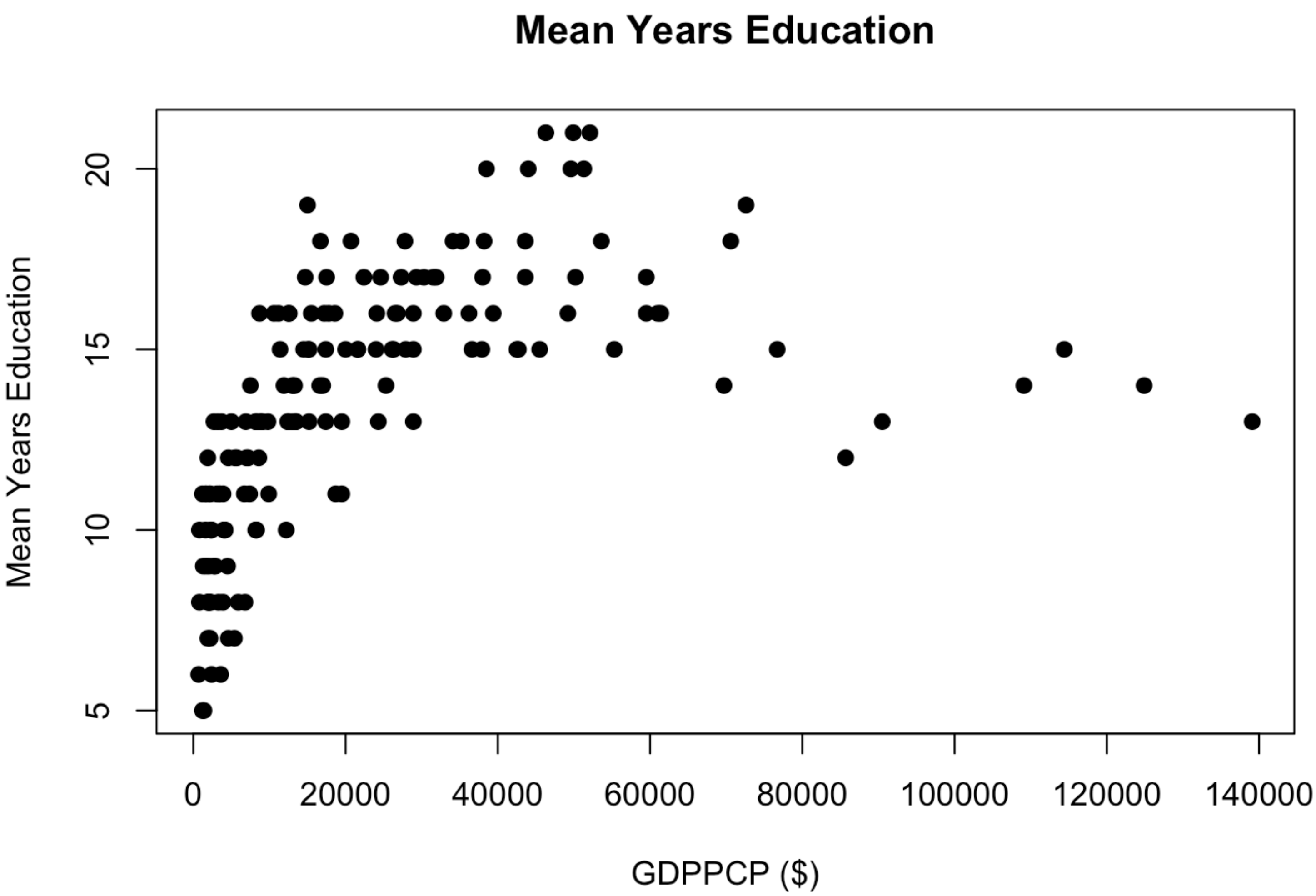
```
# IMPORT DATA
library("readxl")
EduData <- read_excel("Education.xlsx", 1)

# CREATE NEW DATAFRAME WITH WANTED COLUMNS
EduData <- data.frame(EduData$`gdp per capita purchasing power parity usd`,
                      EduData$`school life expectancy total`,
                      EduData$`school life expectancy male`,
                      EduData$`school life expectancy female`)

# CHANGE COLUMN NAMES
colnames(EduData) <- c("GDPPCP",
                       "Total",
                       "Male",
                       "Female")

# CLEAN UP DATA
EduData <- EduData[ grep(0.4242, EduData$GDPPCP, invert = TRUE) , ]
EduData <- EduData[ grep(0.4242, EduData$Total, invert = TRUE) , ]
EduData <- EduData[ grep(0.4242, EduData$Male, invert = TRUE) , ]
EduData <- EduData[ grep(0.4242, EduData$Female, invert = TRUE) , ]

# EXAMINE DATA MEAN YEARS OF EDUCATION
x  <- seq(0, 1500000, 1)
plot(EduData$GDPPCP, EduData$Total, main="Mean Years Education", xlab="GDPPCP ($)"
, ylab="Mean Years Education", pch=19)
```




```
plot(EduData$GDPPCP, EduData$Male, main="Mean Years Education", xlab="GDPPCP ($)",
ylab="Mean Years Education", pch=19)
```



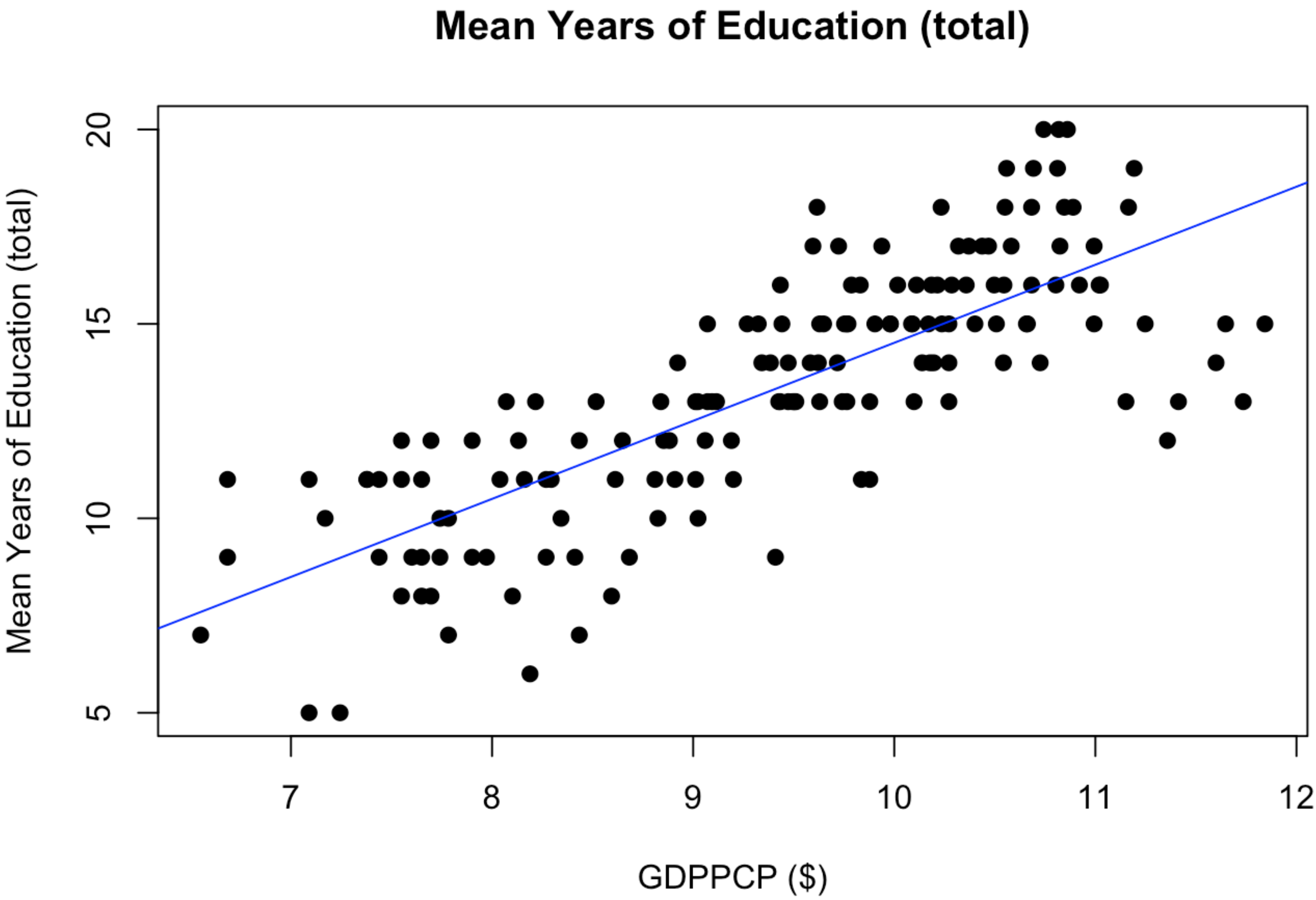

```
plot(EduData$GDPPCP, EduData$Female, main="Mean Years Education", xlab="GDPPCP ($)", ylab="Mean Years Education", pch=19)
```




```
# TRANSFORM DATA FOR EDUCATION
# Here I choose various data transformations to fit the data
# Here I take the log of the GDPPCP (2010) because there are such wide variations
in the dependent variable.
EduData$link_GDPPCP <- log(EduData$GDPPCP)
# This is just a linear relationship; no transformation
EduData$link_EduT <- EduData$Total
EduData$link_EduM <- EduData$Male
EduData$link_EduF <- EduData$Female

# MODEL THE DATA
# Rename the variables with dependent and independent coding
yEduT <- EduData$link_EduT
yEduM <- EduData$link_EduM
yEduF <- EduData$link_EduF
xGDPPCP <- EduData$link_GDPPCP

# REGRESSION ANALYSIS FOR MEAN YEARS OF EDUCATION TOTAL
fit_EduT <- lm(yEduT~xGDPPCP)
summary(fit_EduT) # show results
```

```
##
## Call:
## lm(formula = yEduT ~ xGDPPCP)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.2437 -1.0474  0.1874  1.4152  4.2553
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -5.5605     1.2153  -4.575 9.56e-06 ***
## xGDPPCP       2.0077     0.1279  15.696  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.981 on 158 degrees of freedom
## Multiple R-squared:  0.6093, Adjusted R-squared:  0.6068
## F-statistic: 246.4 on 1 and 158 DF,  p-value: < 2.2e-16
```

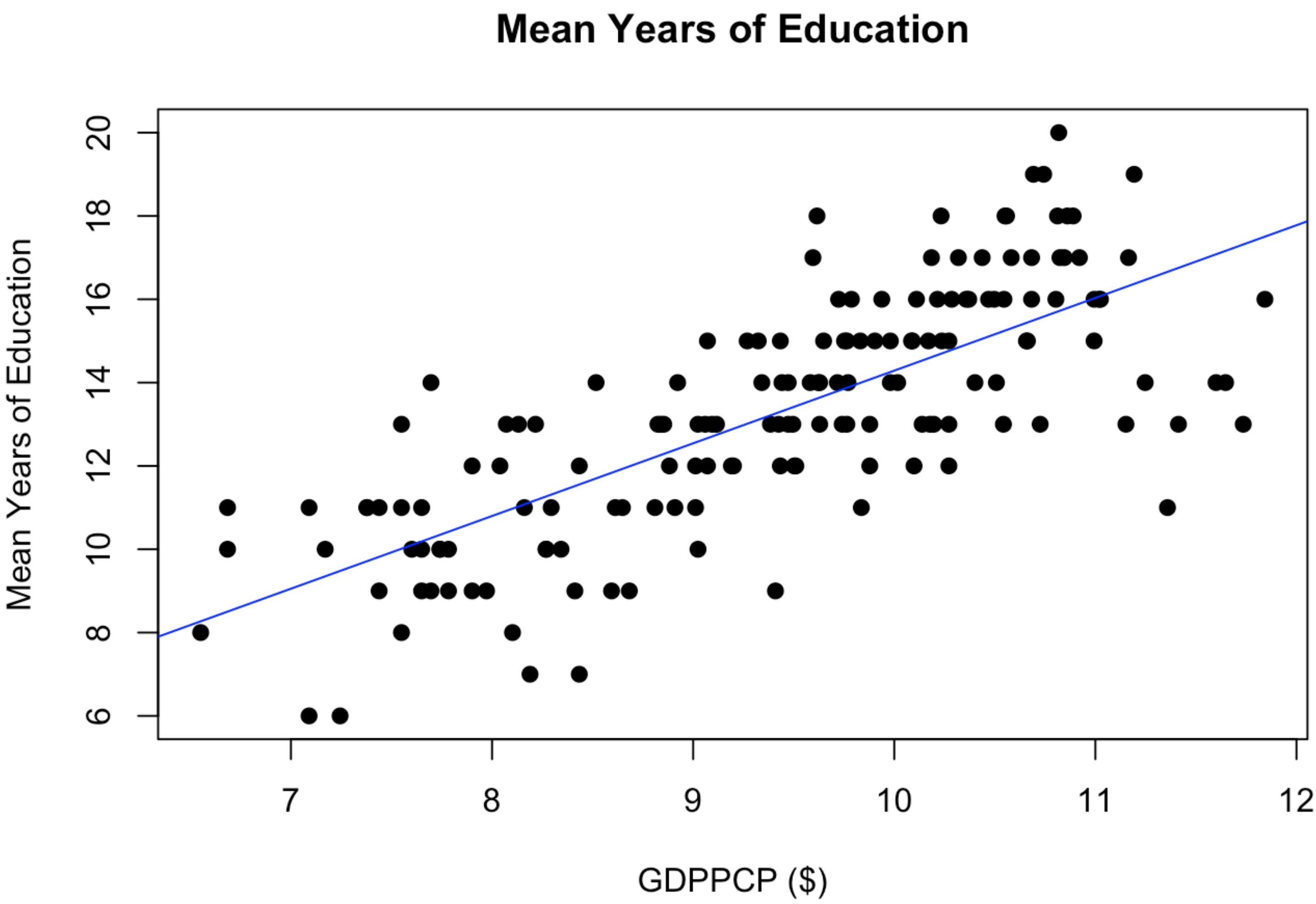
```
r_sqr <-summary(fit_EduT)$r.squared
r_sqr <- round(r_sqr,digits=3)

plot(xGDPPCP, yEduT, main="Mean Years of Education (total)", xlab="GDPPCP ($)", yl
ab="Mean Years of Education (total)", pch=19)
abline(fit_EduT, col="blue") # regression line (y~x)
text(80,3.75,"R Squared is:")
text(80,3.5,r_sqr)
```




```
# REGRESSION ANALYSIS FOR MEAN YEARS OF EDUCATION MALE
fit_EduM <- lm(yEduM~xGDPPCP)
summary(fit_EduM) # show results
```

```
##
## Call:
## lm(formula = yEduM ~ xGDPPCP)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.6607 -1.2679  0.2449  1.2763  4.3810
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -3.164      1.169  -2.706  0.00756 **
## xGDPPCP        1.745      0.123  14.184  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.906 on 158 degrees of freedom
## Multiple R-squared:  0.5601, Adjusted R-squared:  0.5573
## F-statistic: 201.2 on 1 and 158 DF,  p-value: < 2.2e-16
```

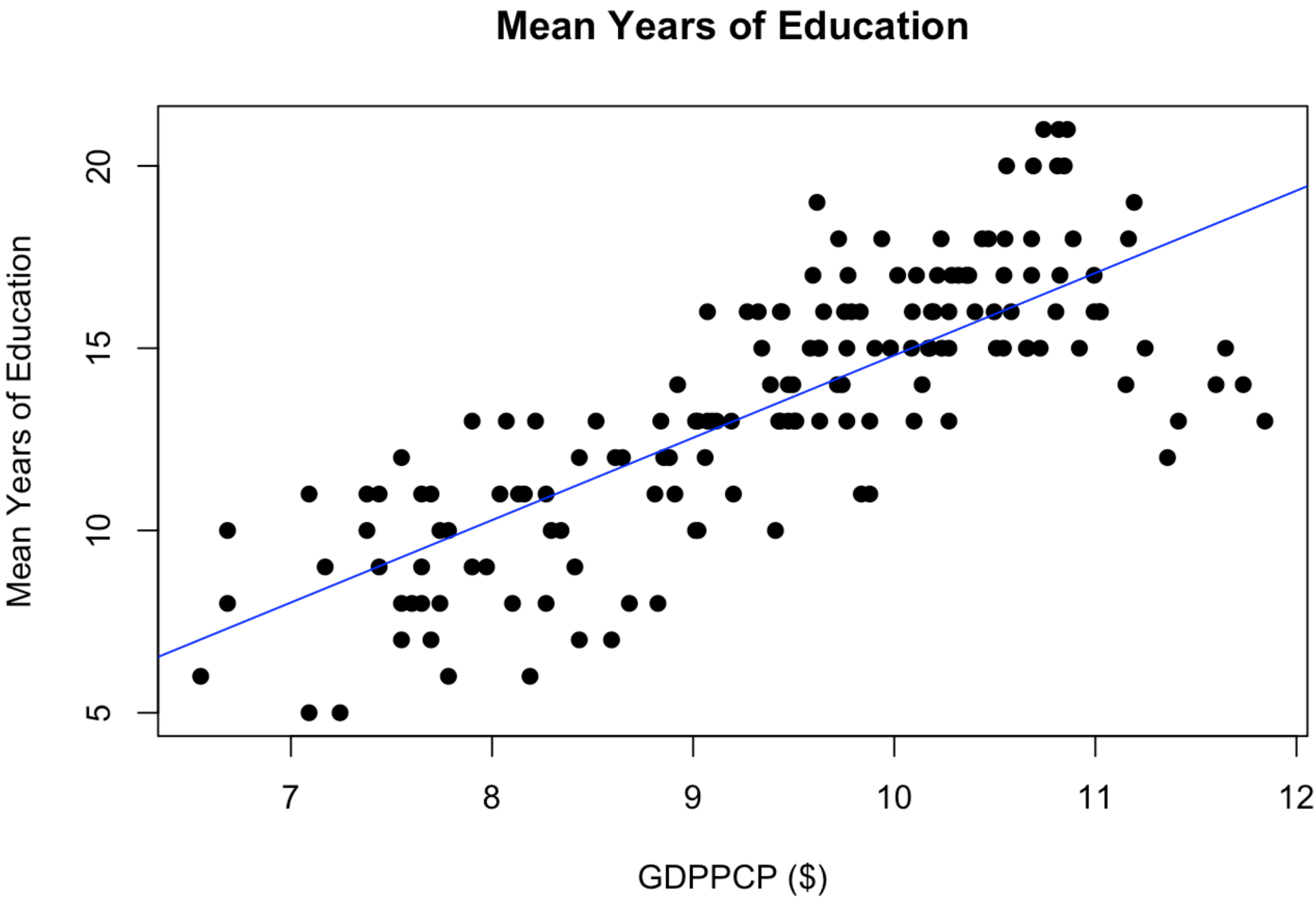
```
r_sqr <-summary(fit_EduM)$r.squared
r_sqr <- round(r_sqr,digits=3)

plot(xGDPPCP, yEduM, main="Mean Years of Education", xlab="GDPPCP ($)", ylab="Mean
Years of Education", pch=19)
abline(fit_EduM, col="blue") # regression line (y~x)
text(80,3.75,"R Squared is:")
text(80,3.5,r_sqr)
```



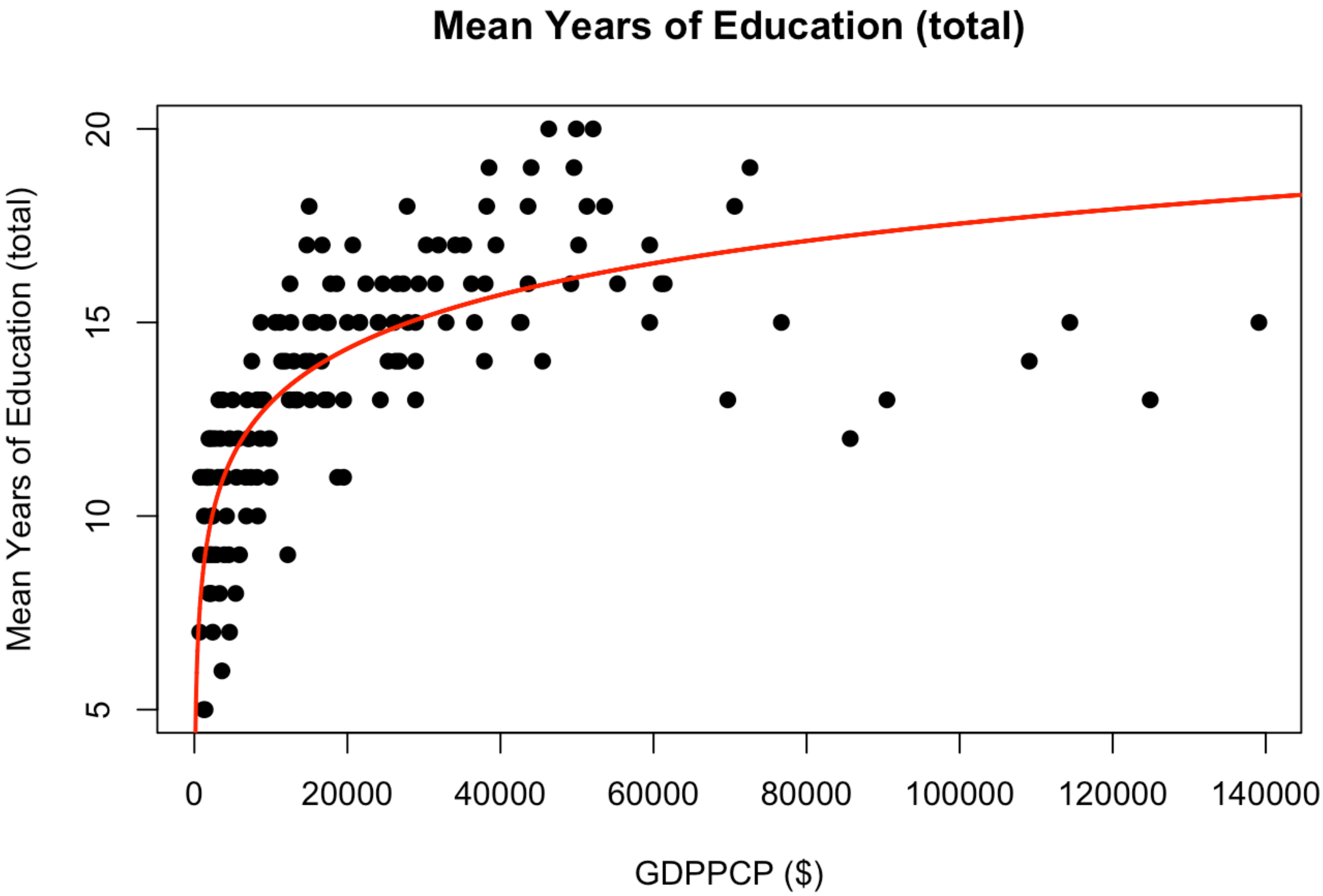

```
# REGRESSION ANALYSIS FOR MEAN YEARS OF EDUCATION FEMALE
fit_EduF <- lm(yEduF~xGDPPCP)
summary(fit_EduF) # show results
```

```
##
## Call:
## lm(formula = yEduF ~ xGDPPCP)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.9709 -1.2343  0.2405  1.5466  5.0627
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -7.7955     1.3692  -5.694 5.91e-08 ***
## xGDPPCP       2.2601     0.1441  15.684  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.232 on 158 degrees of freedom
## Multiple R-squared:  0.6089, Adjusted R-squared:  0.6064
## F-statistic:   246 on 1 and 158 DF,  p-value: < 2.2e-16
```

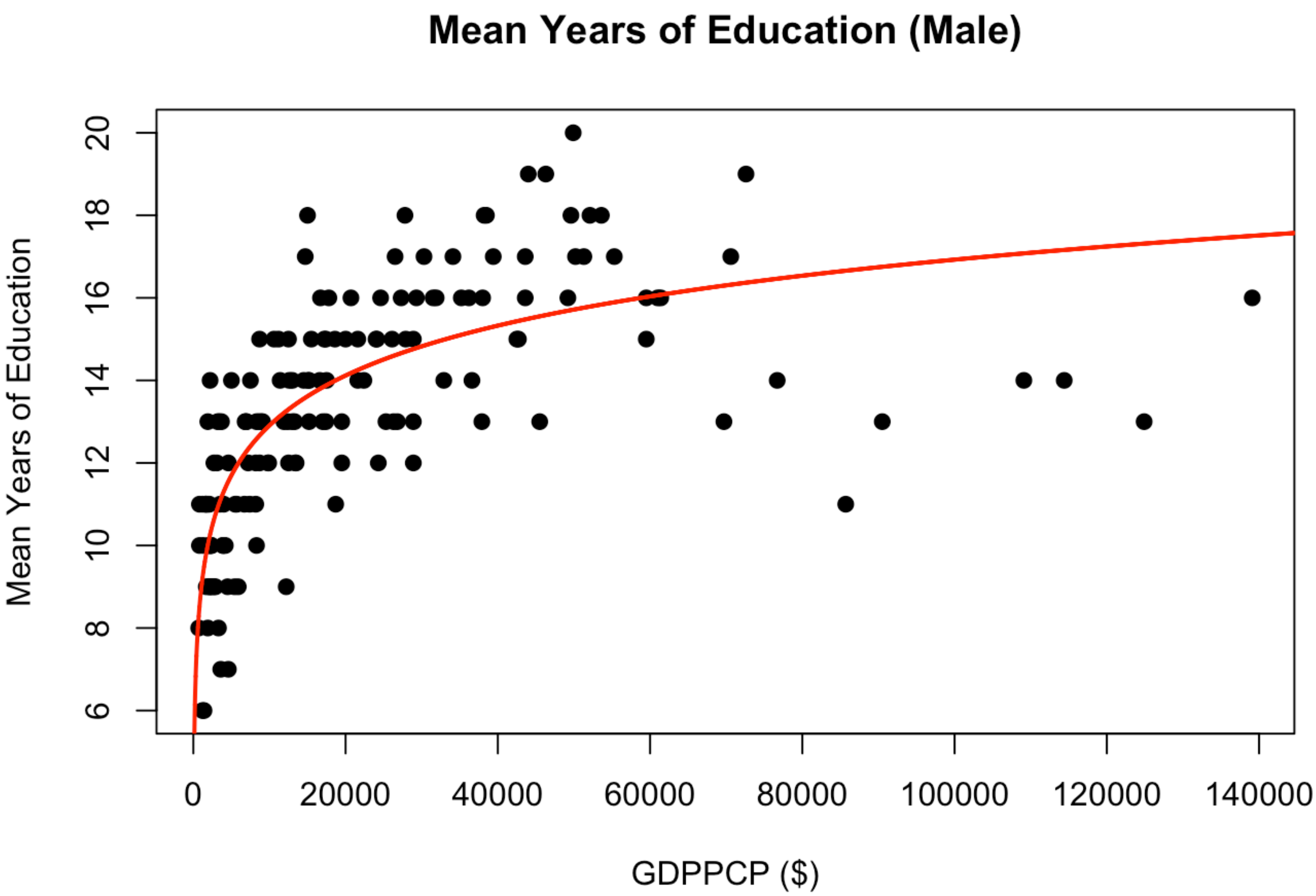
```
r_sqr <-summary(fit_EduF)$r.squared
r_sqr <- round(r_sqr,digits=3)

plot(xGDPPCP, yEduF, main="Mean Years of Education", xlab="GDPPCP ($)", ylab="Mean
Years of Education", pch=19)
abline(fit_EduF, col="blue") # regression line (y~x)
text(80,3.75,"R Squared is:")
text(80,3.5,r_sqr)
```



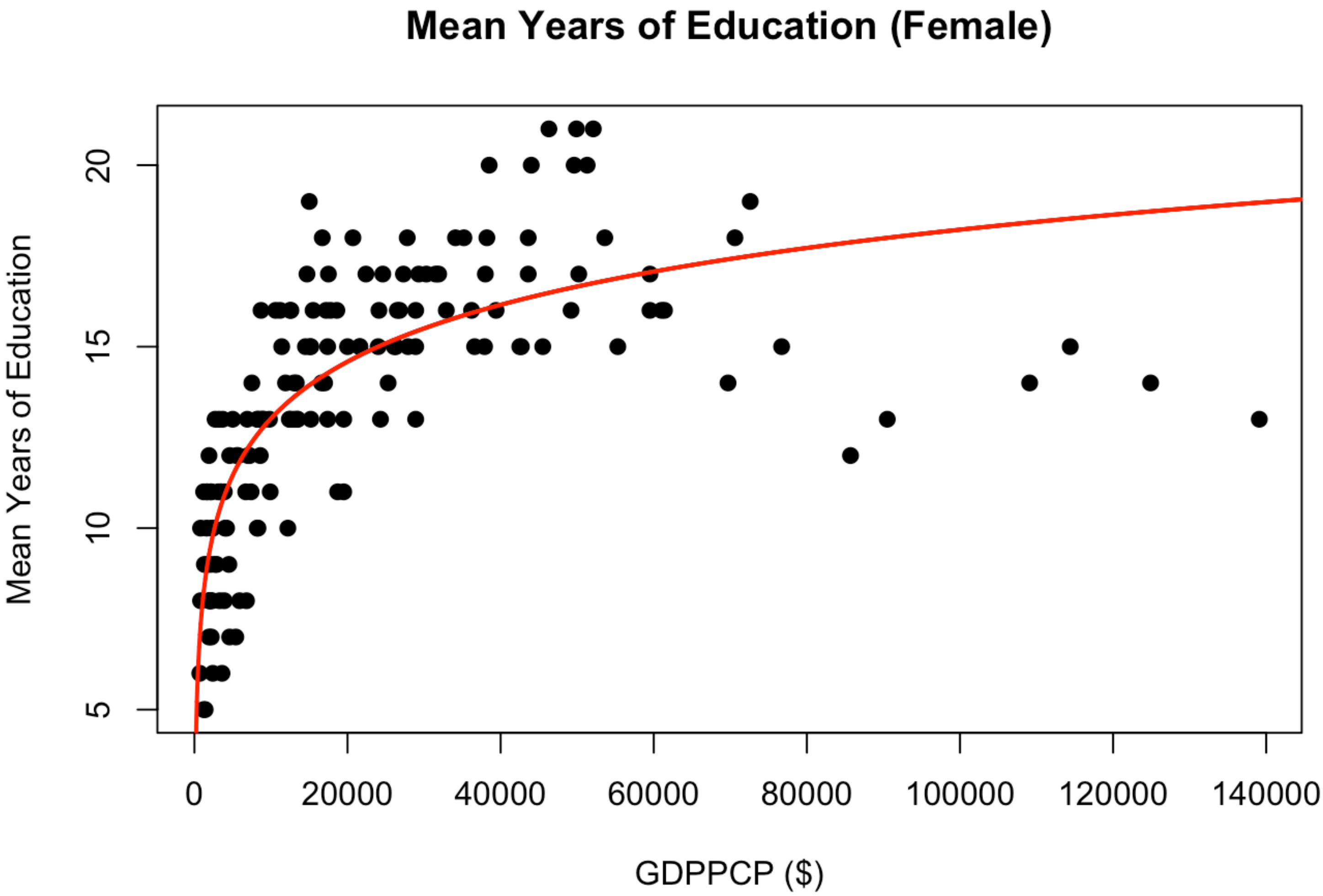

```
# MEAN YEARS OF EDUCATION CURVE TOTAL
x  <- seq(0, 1500000, 1)
plot(EduData$GDPPCP, EduData$Total, main="Mean Years of Education (total)", xlab="
GDPPCP ($)", ylab="Mean Years of Education (total)", pch=19)
lines((2.0077 *log(x))-5.5605, col="red", lwd = 2)
```




```
# MEAN YEARS OF EDUCATION CURVE MALE
x  <- seq(0, 1500000, 1)
plot(EduData$GDPPCP, EduData$Male, main="Mean Years of Education (Male)", xlab="GD
PPCP ($)", ylab="Mean Years of Education", pch=19)
lines((1.745*log(x))-3.164, col="red", lwd = 2)
```




```
# MEAN YEARS OF EDUCATION CURVE FEMALE
x  <- seq(0, 1500000, 1)
plot(EduData$GDPPCP, EduData$Female, main="Mean Years of Education (Female)", xlab
="GDPPCP ($)", ylab="Mean Years of Education", pch=19)
lines((2.2601*log(x))-7.7955, col="red", lwd = 2)
```



Mean Years of Education (Female)
Mean Years of Education
5
10
15
20
0
20000
40000
60000
80000
100000
120000
140000
GDPPCP ($)

B.2. Life Expectancy Male/Female

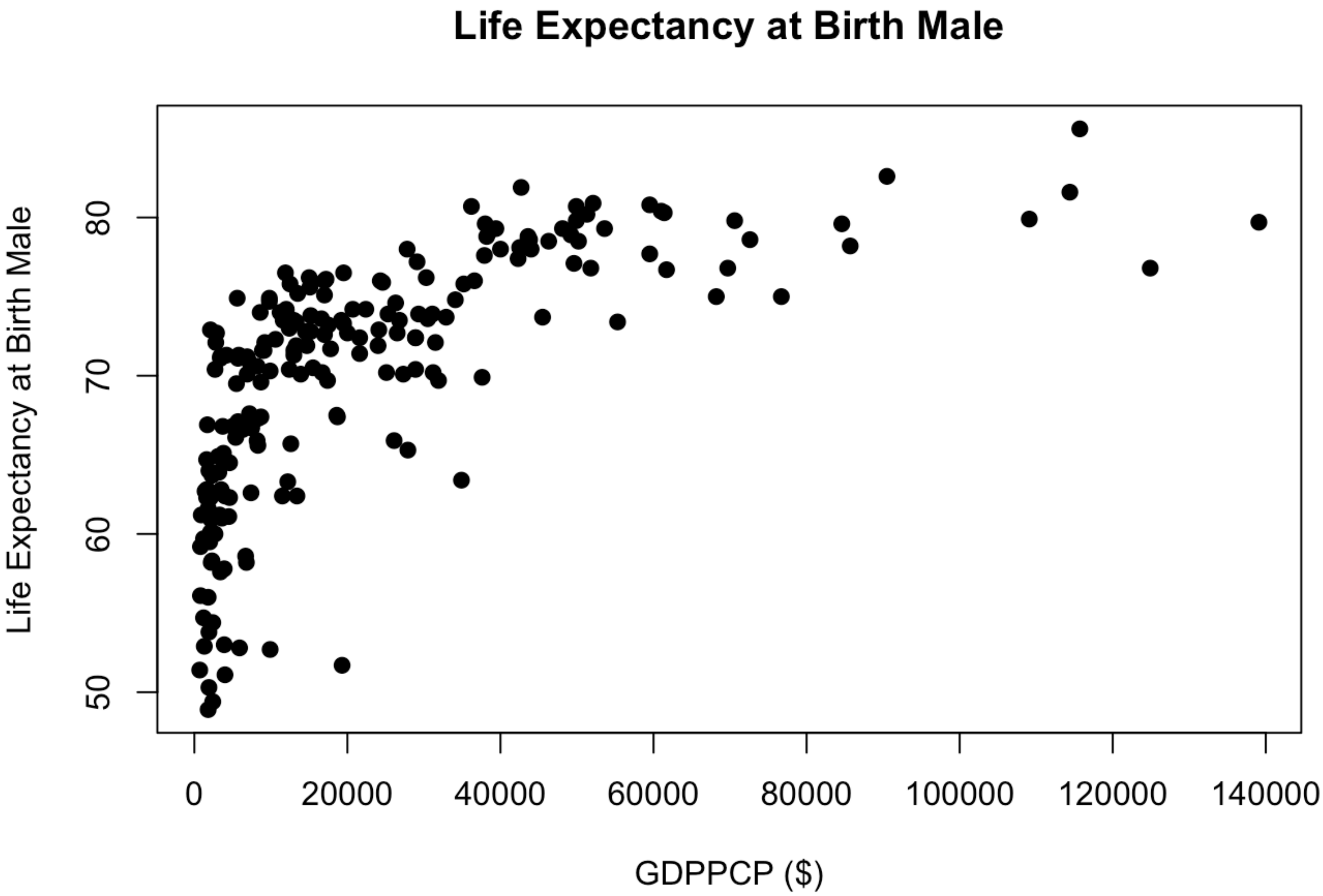
Notebook Thesis Rémon ten Bhömer TU Delft 2018: Regression Curves Life Expectancy Male and Female

```
# IMPORT DATA
library("readxl")
CIAData <- read_excel("RegressionCIA.xlsx", 1)
```

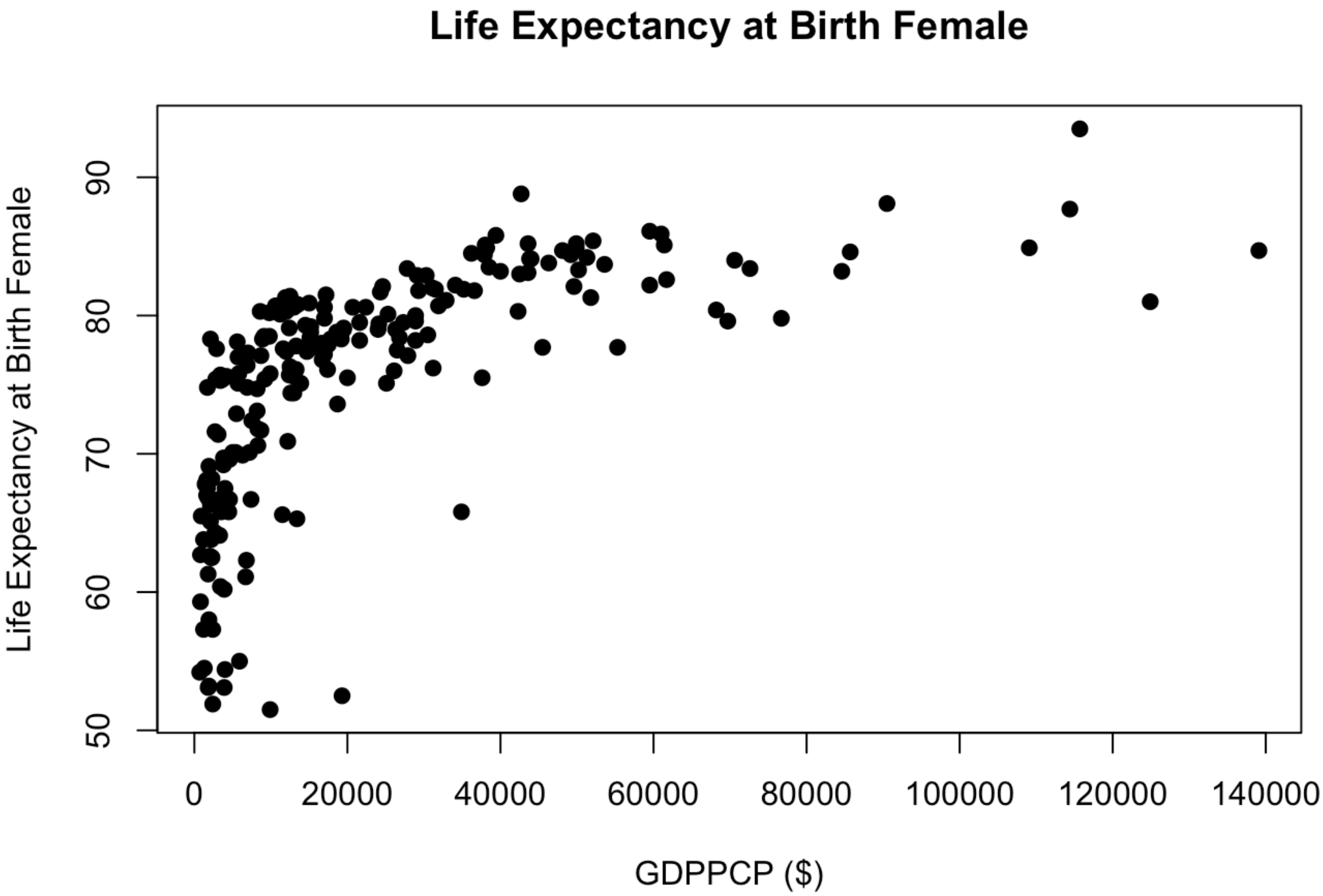
```
# RENAME COLUMNS
colnames(CIAData) <- c("Country","MU5M", "MU5F", "LEM", "LEF", "GDPPCP")
```

```
# CLEAN DATA
CIAData <- CIAData[ grep(0.4242, CIAData$MU5M, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$MU5F, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$LEM, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$LEF, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$GDPPCP, invert = TRUE) , ]
```

```
# EXAMINE LIFE EXPECTANCY
x  <- seq(0, 1500000, 1)
plot(CIAData$GDPPCP, CIAData$LEM, main="Life Expectancy at Birth Male", xlab="GDPP
CP ($)", ylab="Life Expectancy at Birth Male", pch=19)
```




```
x  <- seq(0, 1500000, 1)
plot(CIAData$GDPPCP, CIAData$LEF, main="Life Expectancy at Birth Female", xlab="GD
PPCP ($)", ylab="Life Expectancy at Birth Female", pch=19)
```

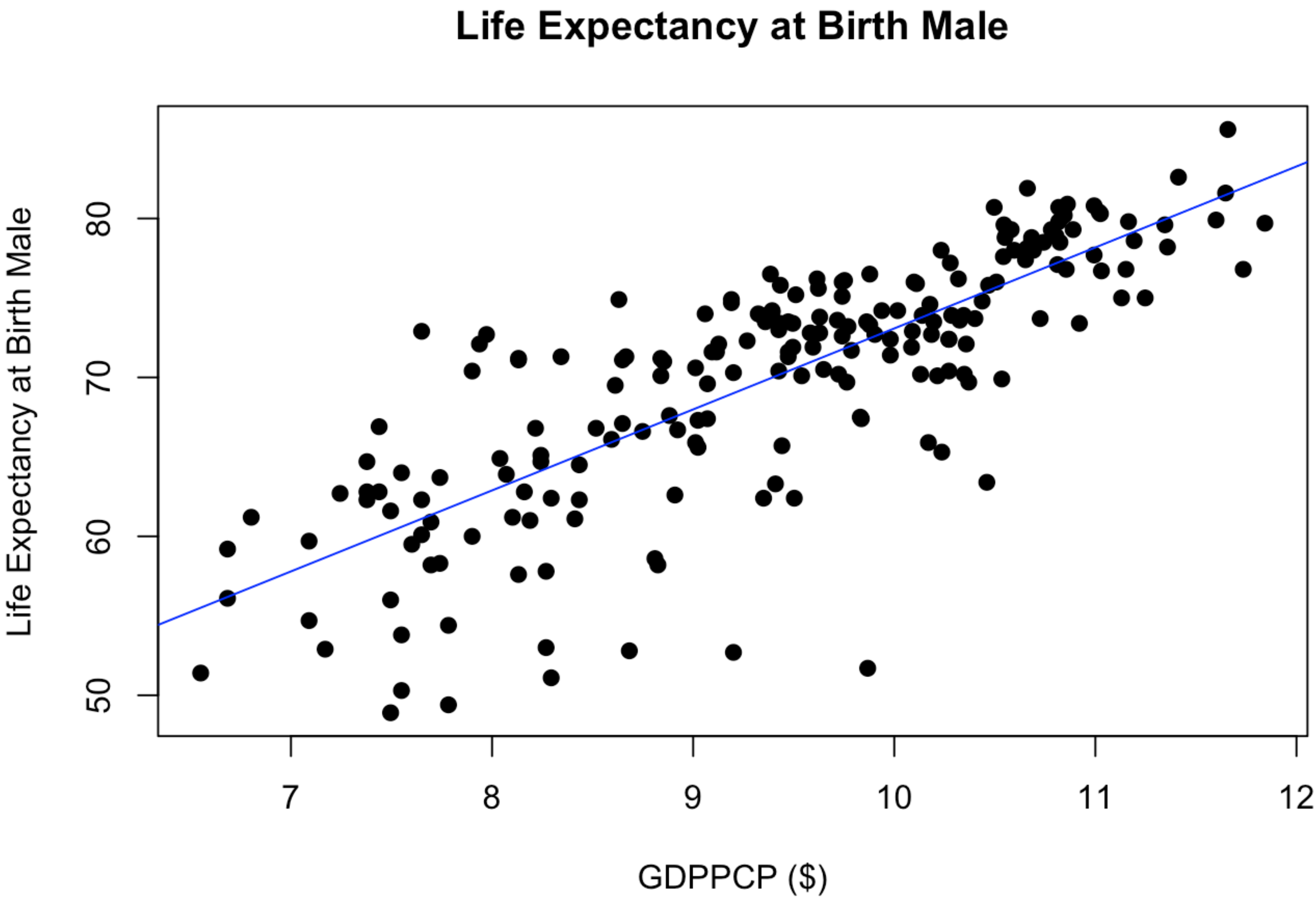



```
# TRANSFORM DATA FOR LIFE EXPECTANCY
# Here I choose various data transformations to fit the data
# Here I take the log of the GDPPCP (2010) because there are such wide variations
in the dependent variable.
CIAData$link_GDPPCP <- log(CIAData$GDPPCP)
# This is just a linear relationship; no transformation
CIAData$link_LEM <- CIAData$LEM
CIAData$link_LEF <- CIAData$LEF

# MODEL THE DATA
# Rename the variables with dependent and independent coding
yLEM <- CIAData$link_LEM
yLEF <- CIAData$link_LEF
xGDPPCP <- CIAData$link_GDPPCP
```

```
# REGRESSION ANALYSIS FOR LIFE EXPECTANCY MALE
fit_LEM <- lm(yLEM~xGDPPCP)
summary(fit_LEM) # show results
```

```
##
## Call:
## lm(formula = yLEM ~ xGDPPCP)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.7127  -2.1102   0.9014   2.9359  11.7936
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  22.1151     2.5618   8.633 1.67e-15 ***
## xGDPPCP       5.0971     0.2696  18.909  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.712 on 205 degrees of freedom
## Multiple R-squared:  0.6356, Adjusted R-squared:  0.6338
## F-statistic: 357.6 on 1 and 205 DF,  p-value: < 2.2e-16
```

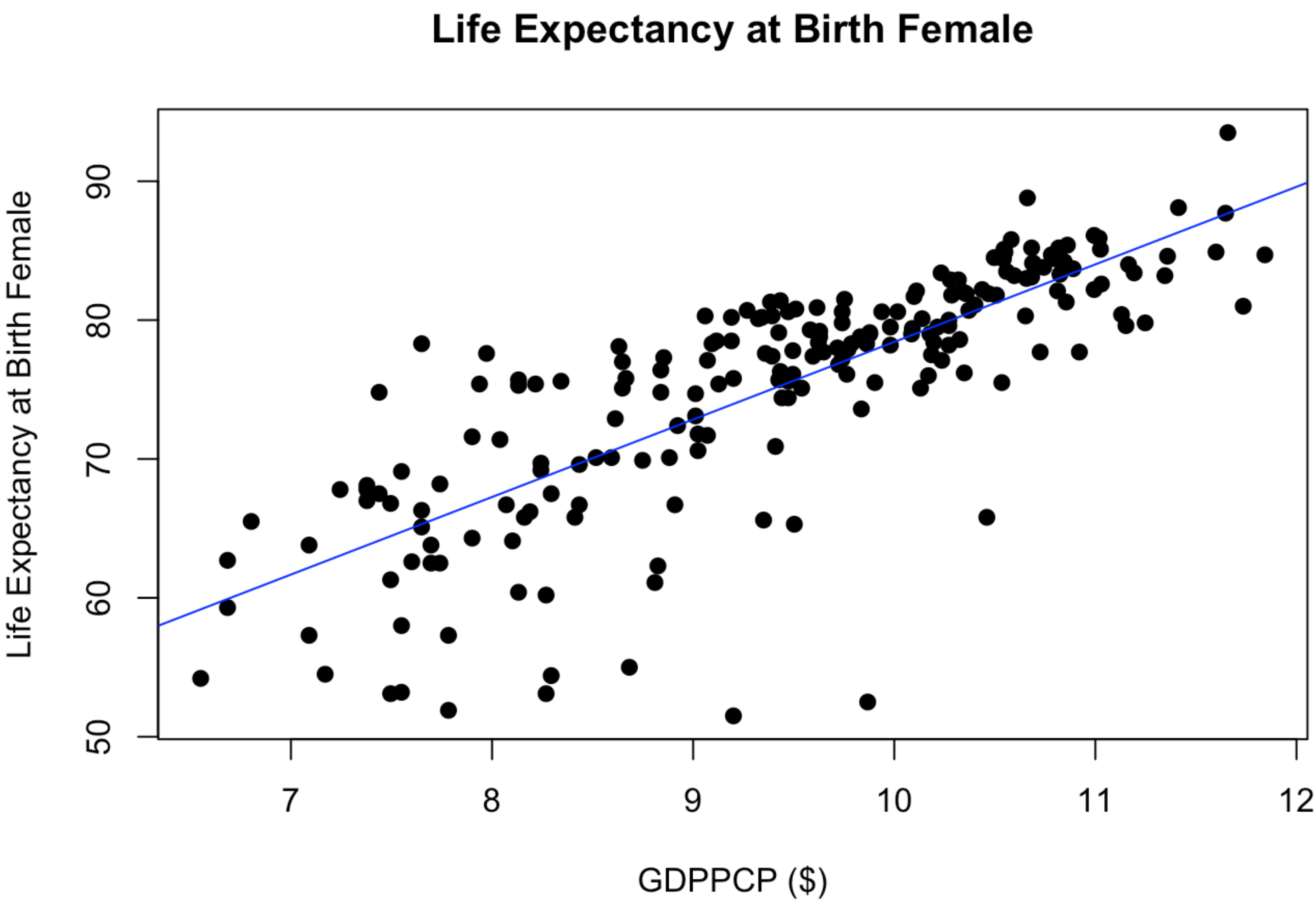
```
r_sqr <-summary(fit_LEM)$r.squared
r_sqr <- round(r_sqr,digits=3)

plot(xGDPPCP, yLEM, main="Life Expectancy at Birth Male", xlab="GDPPCP ($)", ylab=
"Life Expectancy at Birth Male", pch=19)
abline(fit_LEM, col="blue") # regression line (y~x)
text(80,3.75,"R Squared is:")
text(80,3.5,r_sqr)
```



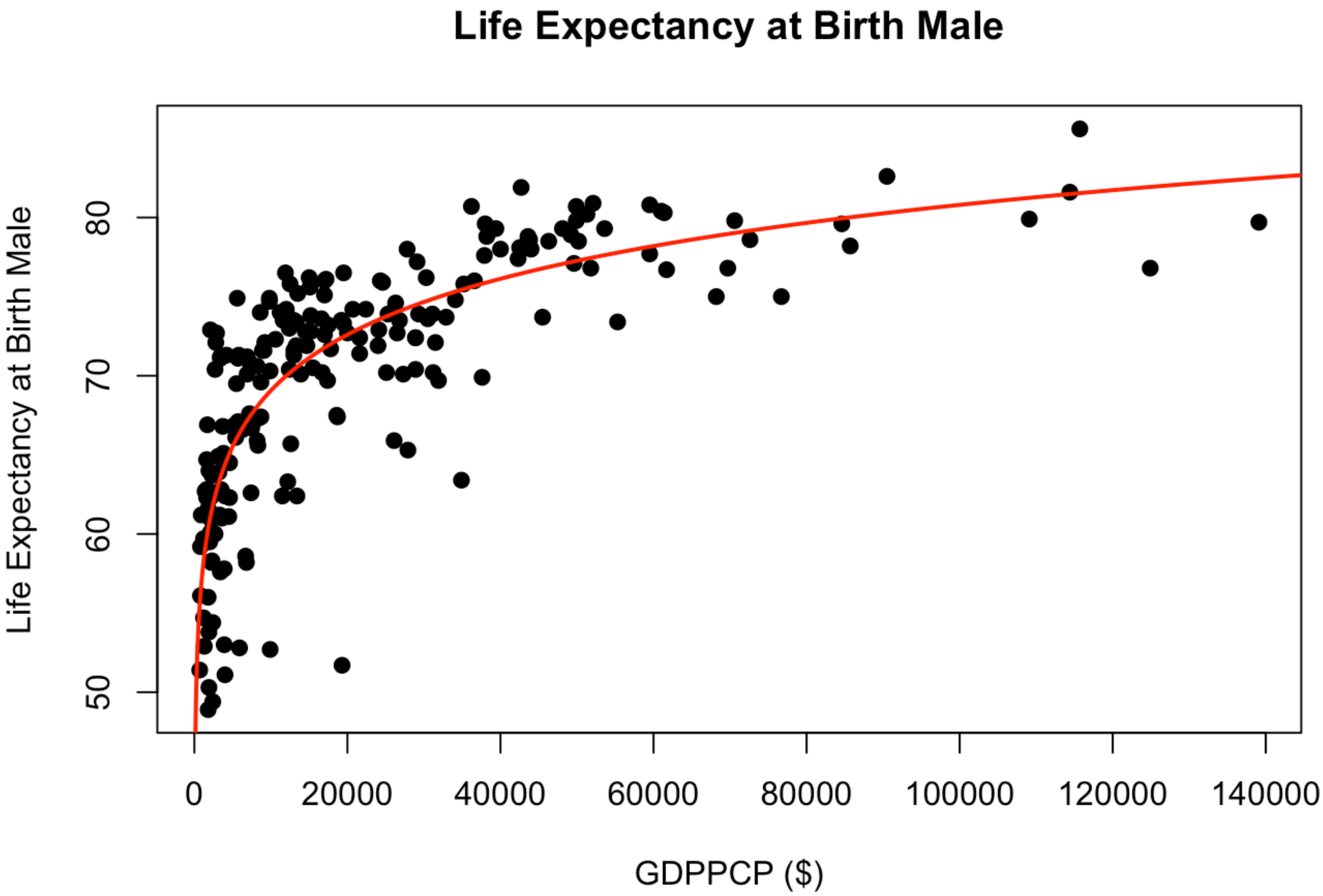

```
# REGRESSION ANALYSIS FOR LIFE EXPECTANCY FEMALE
fit_LEF <- lm(yLEF~xGDPPCP)
summary(fit_LEF) # show results
```

```
##
## Call:
## lm(formula = yLEF ~ xGDPPCP)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -25.1941  -1.8449   0.8972   2.9248  12.9941
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  22.5833     2.8725   7.862 2.12e-13 ***
## xGDPPCP       5.5849     0.3023  18.478  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.283 on 205 degrees of freedom
## Multiple R-squared:  0.6248, Adjusted R-squared:  0.623
## F-statistic: 341.4 on 1 and 205 DF,  p-value: < 2.2e-16
```

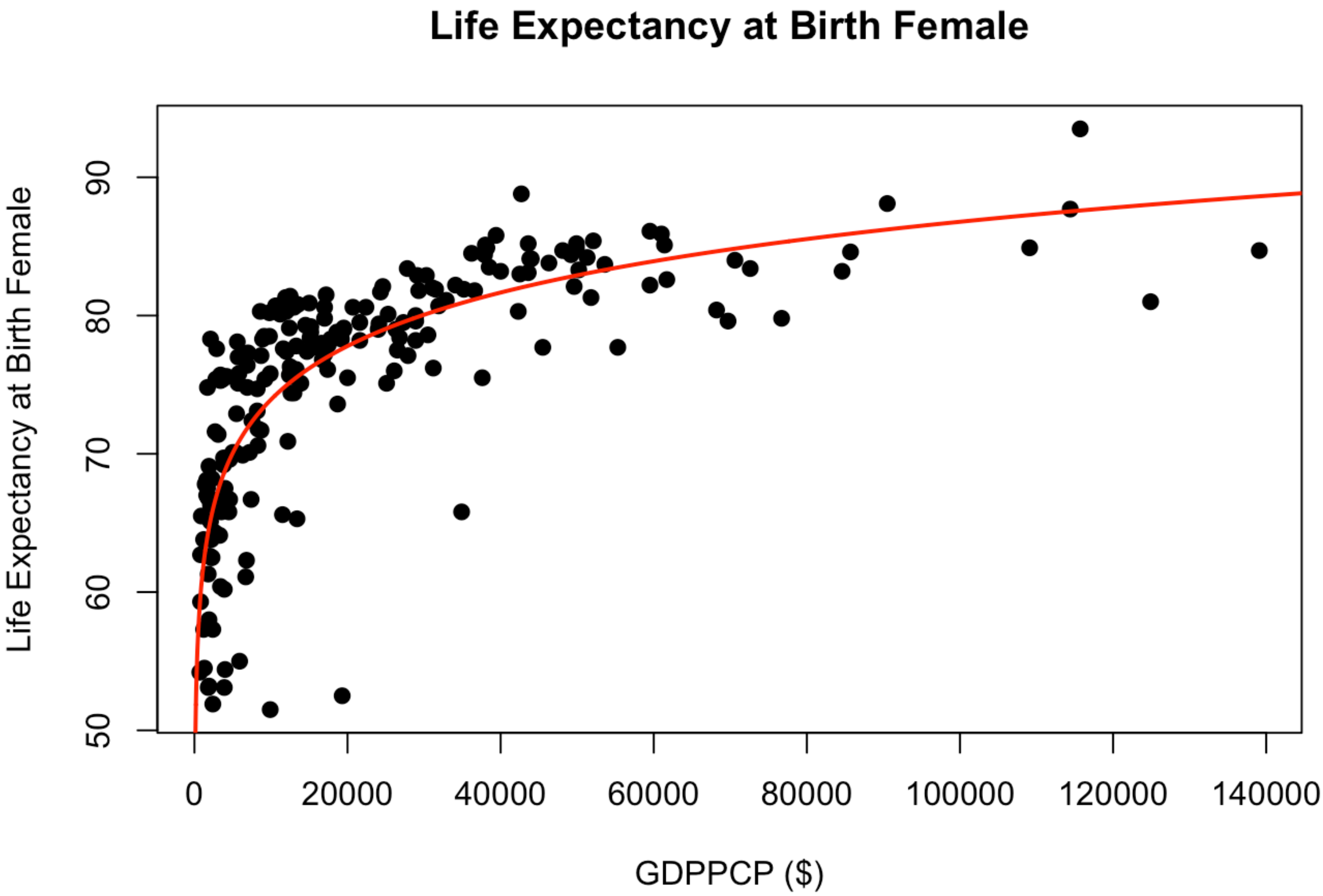
```
r_sqr <-summary(fit_LEF)$r.squared
r_sqr <- round(r_sqr,digits=3)

plot(xGDPPCP, yLEF, main="Life Expectancy at Birth Female", xlab="GDPPCP ($)", yla
b="Life Expectancy at Birth Female", pch=19)
abline(fit_LEF, col="blue") # regression line (y~x)
text(80,3.75,"R Squared is:")
text(80,3.5,r_sqr)
```




```
# LIFE EXPECTANCY REGRESSION CURVE MALE
x  <- seq(0,90,0.01)
x_hat <- exp(x)
y <- 5.0971*x+22.1151
y_hat <- exp(y)
plot(CIAData$GDPPCP, CIAData$LEM, main="Life Expectancy at Birth Male", xlab="GDPP
CP ($)", ylab="Life Expectancy at Birth Male", pch=19)
points(x_hat,y, type='l', col="red", lwd = 2)
```




```
# LIFE EXPECTANCY REGRESSION CURVE FEMALE
x  <- seq(0,90,0.01)
x_hat <- exp(x)
y <- 5.5849*x+22.4833
y_hat <- exp(y)
plot(CIAData$GDPPCP, CIAData$LEF, main="Life Expectancy at Birth Female", xlab="GD
PPCP ($)", ylab="Life Expectancy at Birth Female", pch=19)
points(x_hat,y, type='l', col="red", lwd = 2)
```



Life Expectancy at Birth Female
Life Expectancy at Birth Female
50
60
70
80
90
0
20000
40000
60000
80000
100000
120000
140000
GDPPCP ($)

B.3. Birth Rate

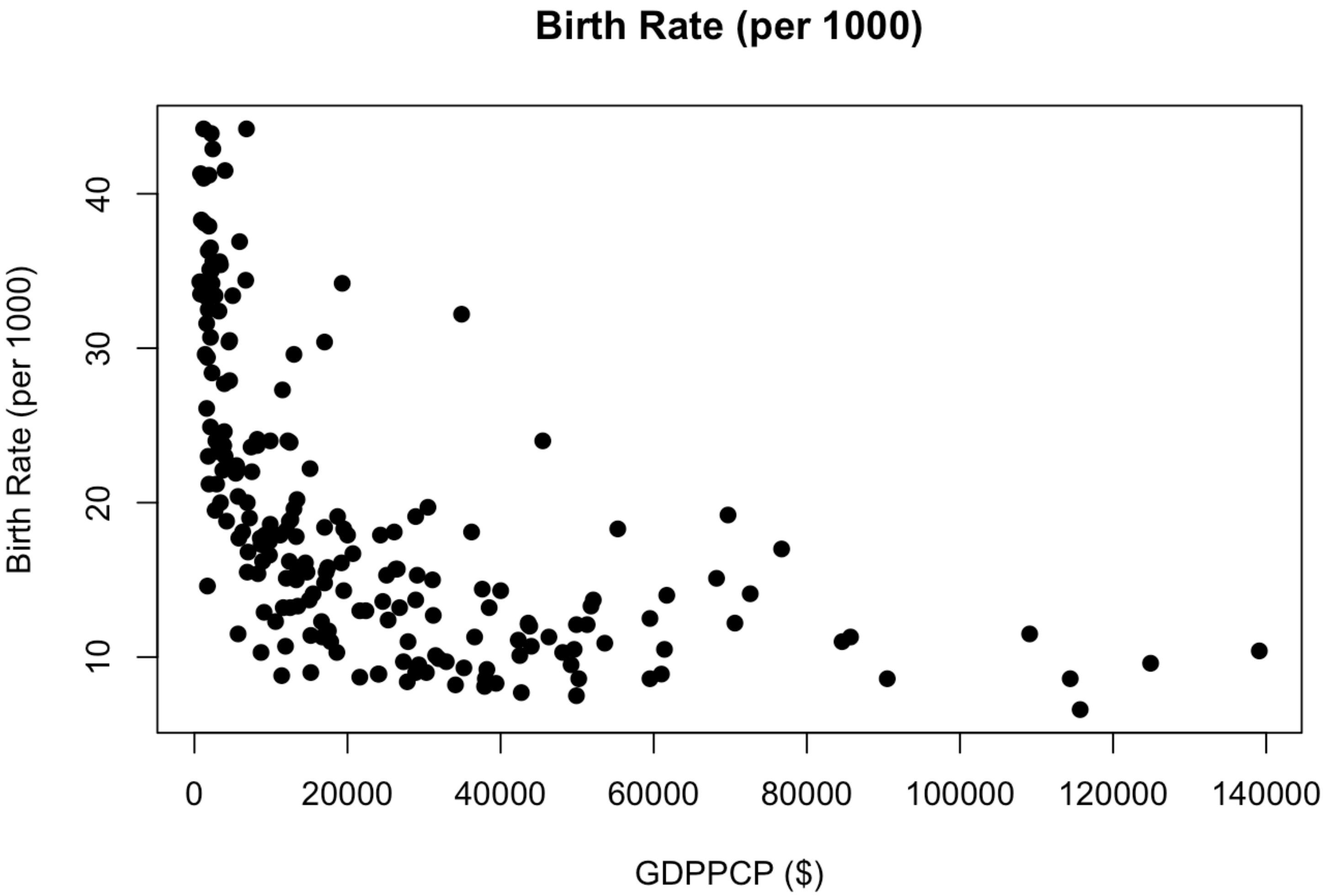
Notebook Thesis Rémon ten Bhömer TU Delft 2018: Regression Curve Birth Rate

```
library("readxl")
CIAData <- read_excel("RegressionCIA.xlsx", 1)
```

```
# RENAME THE COLUMN NAMES
colnames(CIAData) <- c("Country","MU5M", "MU5F", "LEM", "LEF", "GDPPCP", "BR")
```

```
# CLEAN THE DATA
# DELETE FROM EVERY COLUMN THE ROWS WITH 0.4242 AS VALUE
CIAData <- CIAData[ grep(0.4242, CIAData$MU5M, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$MU5F, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$LEM, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$LEF, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$GDPPCP, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$BR, invert = TRUE) , ]
```

```
# EXAMINE BIRTH RATE
x  <- seq(0, 1500000, 1)
plot(CIAData$GDPPCP, CIAData$BR, main="Birth Rate (per 1000)", xlab="GDPPCP ($)",
ylab="Birth Rate (per 1000)", pch=19)
```

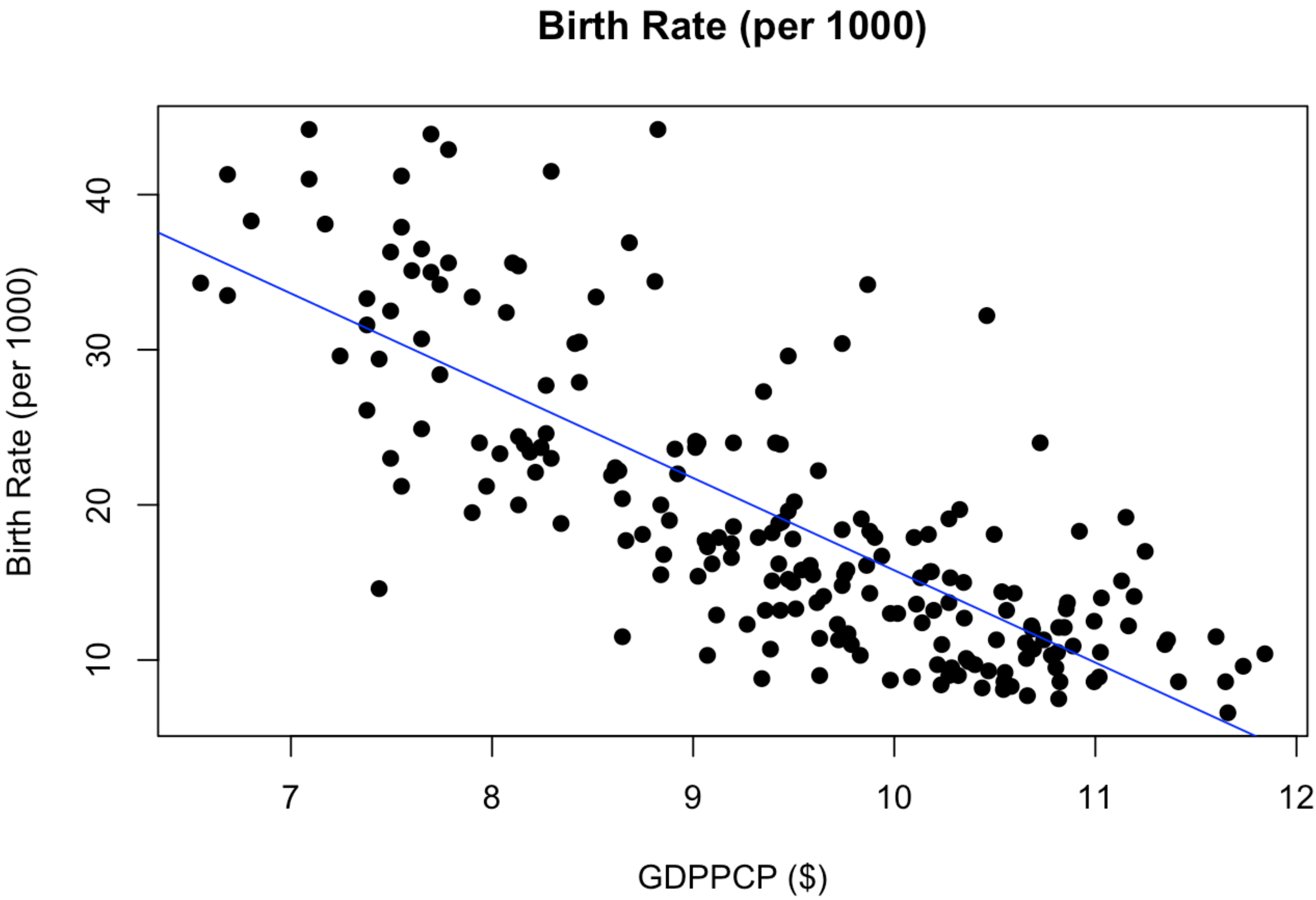



```
# TRANSFORM DATA FOR BIRTH RATE
# Here I choose various data transformations to fit the data
# Here I take the log of the GDPPCP ($) because there are such wide variations in
the dependent variable.
CIAData$link_GDPPCP <- log(CIAData$GDPPCP)
# This is just a linear relationship; no transformation
CIAData$link_BR <- CIAData$BR

# MODEL THE DATA
# Rename the variables with dependent and independent coding
yBR <- CIAData$link_BR
xGDPPCP <- CIAData$link_GDPPCP
```

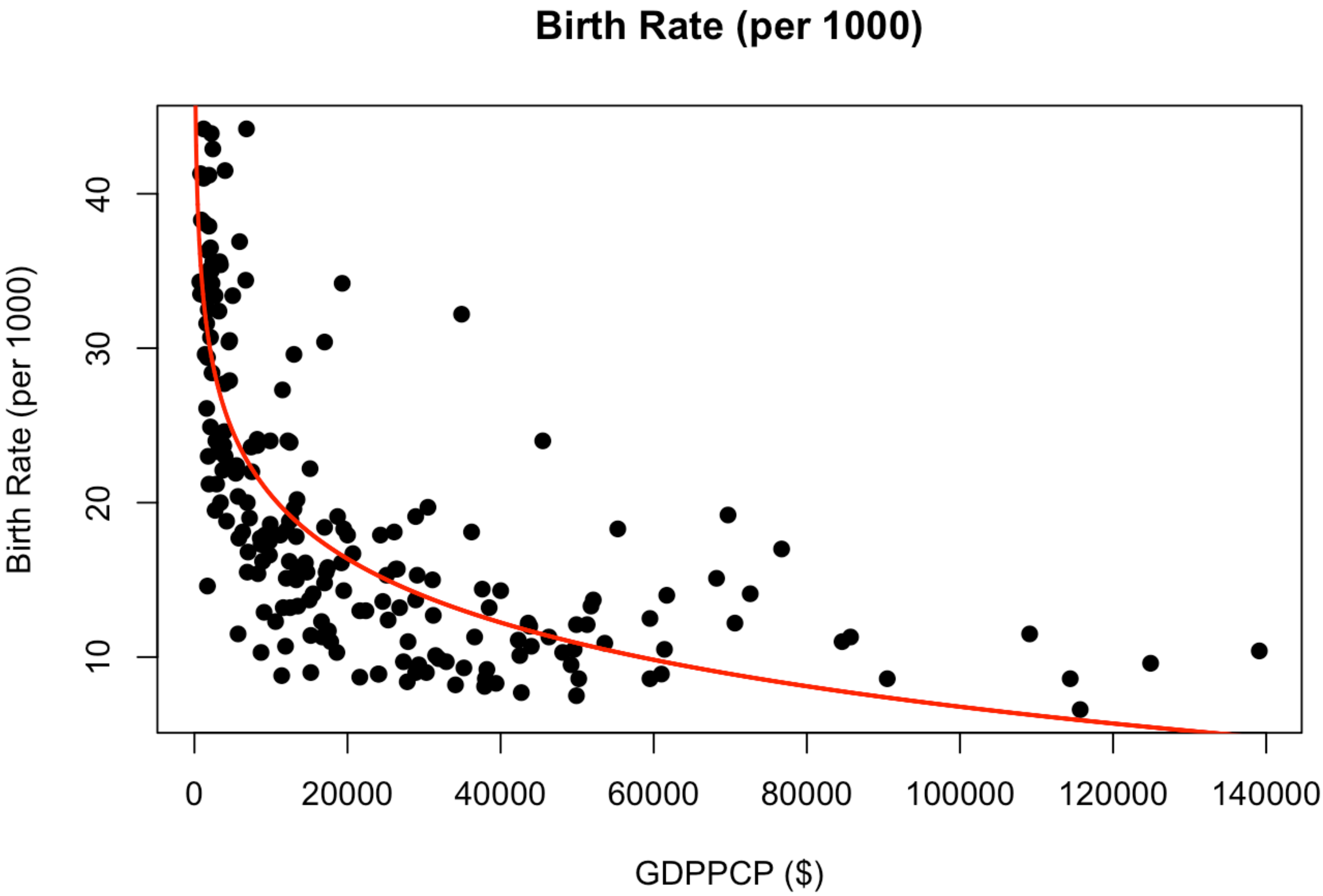
```
# REGRESSION ANALYSIS FOR BIRTH RATE
fit_BR <- lm(yBR~xGDPPCP)
summary(fit_BR) # show results
```

```
##
## Call:
## lm(formula = yBR ~ xGDPPCP)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.4302  -3.7839  -0.9893   3.1277  21.4188
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  75.2920     3.1807   23.67   <2e-16 ***
## xGDPPCP      -5.9505     0.3347  -17.78   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.85 on 205 degrees of freedom
## Multiple R-squared:  0.6066, Adjusted R-squared:  0.6047
## F-statistic: 316.1 on 1 and 205 DF,  p-value: < 2.2e-16
```

```
r_sqr <-summary(fit_BR)$r.squared
r_sqr <- round(r_sqr,digits=3)

plot(xGDPPCP, yBR, main="Birth Rate (per 1000)", xlab="GDPPCP ($)", ylab="Birth Ra
te (per 1000)", pch=19)
abline(fit_BR, col="blue") # regression line (y~x)
text(80,3.75,"R Squared is:")
text(80,3.5,r_sqr)
```




```
# BIRTH RATE REGRESSION CURVE
x  <- seq(0, 1500000, 1)
plot(CIAData$GDPPCP, CIAData$BR, main="Birth Rate (per 1000)", xlab="GDPPCP ($)",
ylab="Birth Rate (per 1000)", pch=19)
lines((-5.9505*log(x))+75.2920, col="red", lwd = 2)
```



Birth Rate (per 1000)
Birth Rate (per 1000)
40
30
20
10
0
20000
40000
60000
80000
100000
120000
140000
GDPPCP ($)

B.4. Infant Mortality Rate Under 5 Male/Female

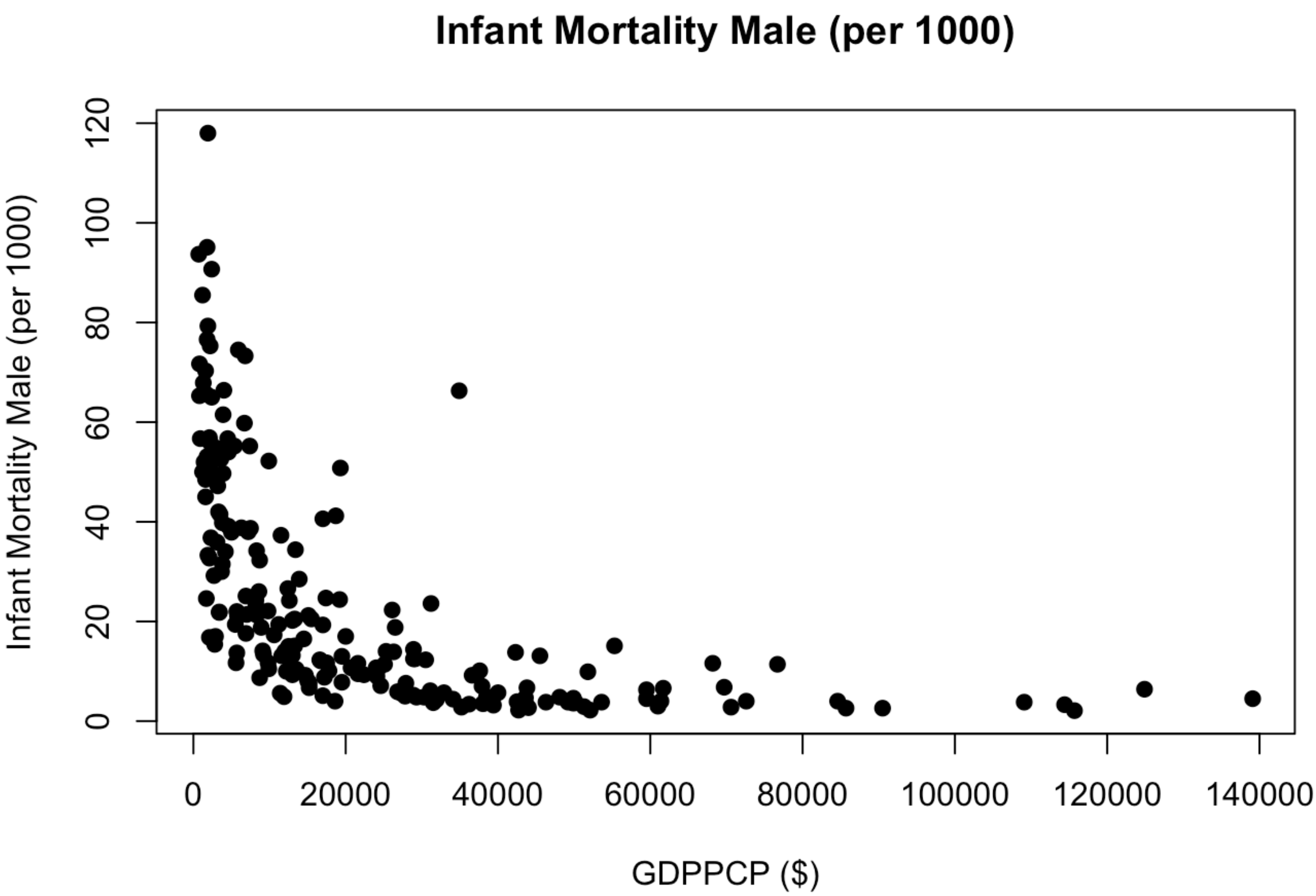
Notebook Thesis Rémon ten Bhömer TU Delft 2018: Regression Curves Infant Mortality Under 5

```
library("readxl")
CIAData <- read_excel("RegressionCIA.xlsx", 1)
```

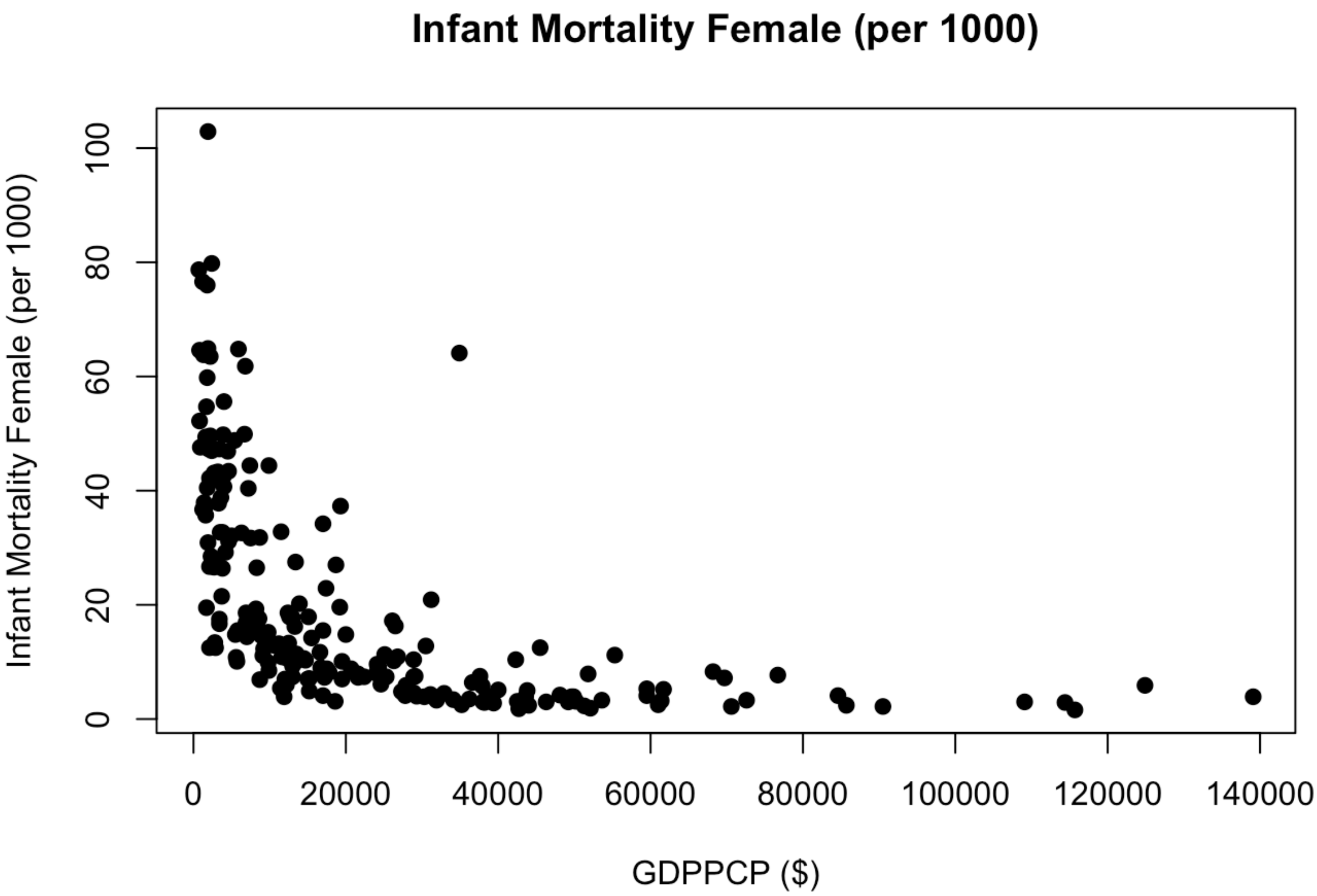
```
colnames(CIAData) <- c("Country","MU5M", "MU5F", "LEM", "LEF", "GDPPCP")
```

```
CIAData <- CIAData[ grep(0.4242, CIAData$MU5M, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$MU5F, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$LEM, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$LEF, invert = TRUE) , ]
CIAData <- CIAData[ grep(0.4242, CIAData$GDPPCP, invert = TRUE) , ]
```

```
# EXAMINE LIFE EXPECTANCY
x  <- seq(0, 1500000, 1)
plot(CIAData$GDPPCP, CIAData$MU5M, main="Infant Mortality Male (per 1000)", xlab="
GDPPCP ($)", ylab="Infant Mortality Male (per 1000)", pch=19)
```



```
x  <- seq(0, 1500000, 1)
plot(CIAData$GDPPCP, CIAData$MU5F, main="Infant Mortality Female (per 1000)", xlab
="GDPPCP ($)", ylab="Infant Mortality Female (per 1000)", pch=19)
```

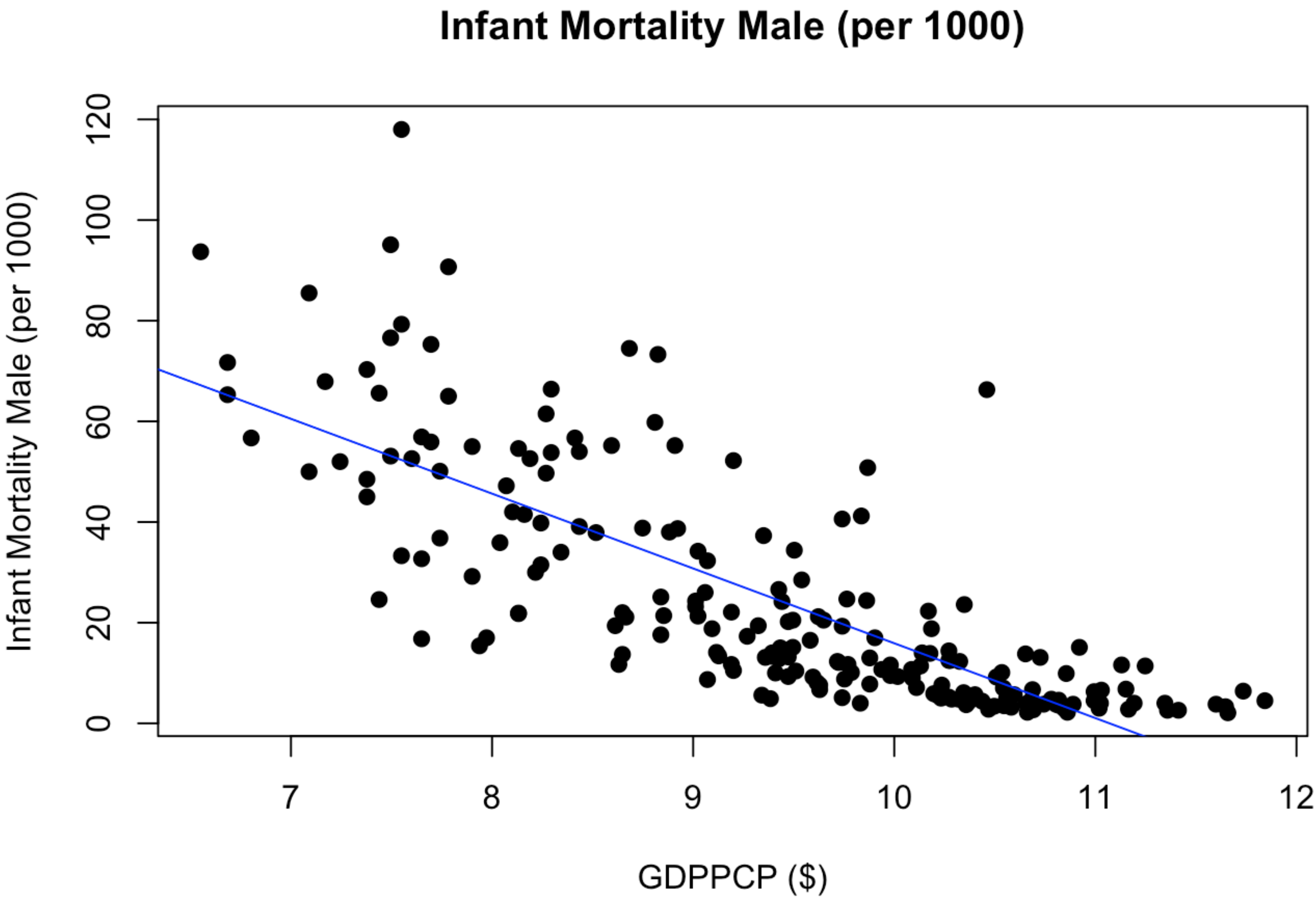



```
# TRANSFORM DATA FOR LIFE EXPECTANCY
# Here I choose various data transformations to fit the data
# Here I take the log of the GDPPCP (2010) because there are such wide variations
in the dependent variable.
CIAData$link_GDPPCP <- log(CIAData$GDPPCP)
# This is just a linear relationship; no transformation
CIAData$link_MU5M <- CIAData$MU5M
CIAData$link_MU5F <- CIAData$MU5F

# MODEL THE DATA
# Rename the variables with dependent and independent coding
yMU5M <- CIAData$link_MU5M
yMU5F <- CIAData$link_MU5F
xGDPPCP <- CIAData$link_GDPPCP
```

```
# REGRESSION ANALYSIS FOR LIFE EXPECTANCY MALE
fit_MU5M <- lm(yMU5M~xGDPPCP)
summary(fit_MU5M) # show results
```

```
##
## Call:
## lm(formula = yMU5M ~ xGDPPCP)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -34.074  -8.202  -1.344   6.456  65.637
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 164.6484     7.8889   20.87   <2e-16 ***
## xGDPPCP     -14.8730     0.8301  -17.92   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.51 on 205 degrees of freedom
## Multiple R-squared:  0.6103, Adjusted R-squared:  0.6084
## F-statistic:   321 on 1 and 205 DF,  p-value: < 2.2e-16
```

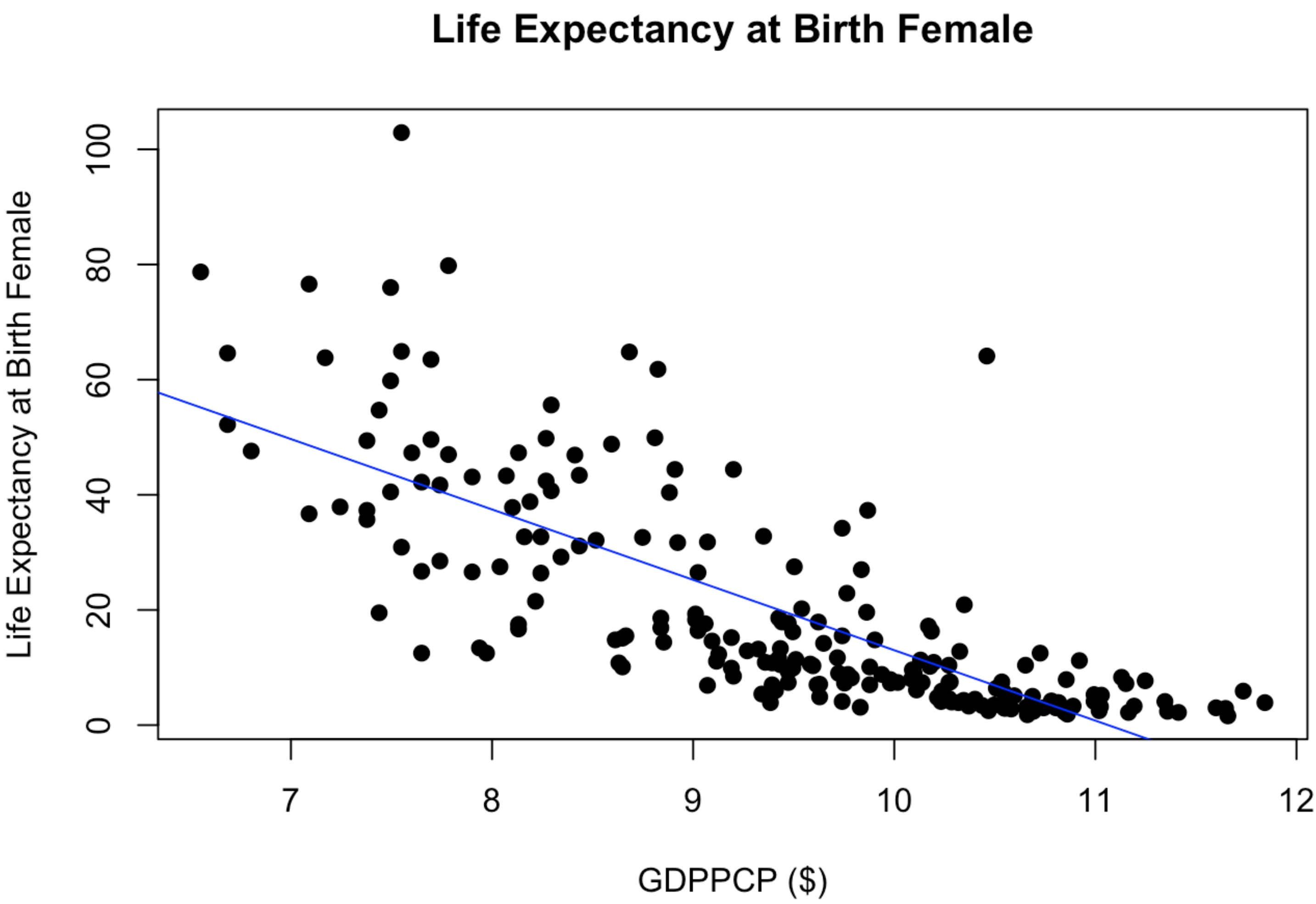
```
r_sqr <-summary(fit_MU5M)$r.squared
r_sqr <- round(r_sqr,digits=3)

plot(xGDPPCP, yMU5M, main="Infant Mortality Male (per 1000)", xlab="GDPPCP ($)", y
lab="Infant Mortality Male (per 1000)", pch=19)
abline(fit_MU5M, col="blue") # regression line (y~x)
text(80,3.75,"R Squared is:")
text(80,3.5,r_sqr)
```



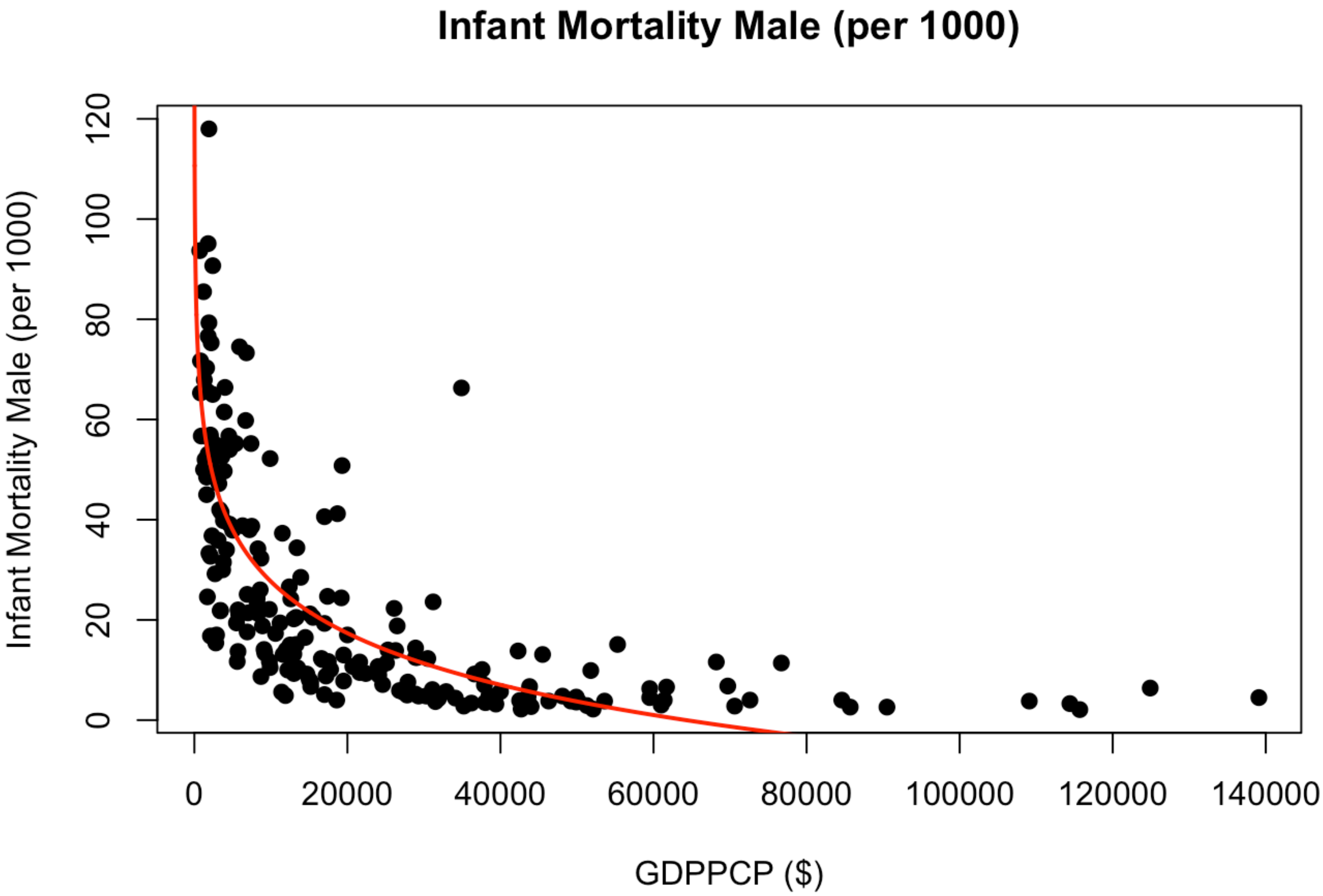

```
# REGRESSION ANALYSIS FOR LIFE EXPECTANCY FEMALE
fit_MU5F <- lm(yMU5F~xGDPPCP)
summary(fit_MU5F) # show results
```

```
##
## Call:
## lm(formula = yMU5F ~ xGDPPCP)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.255  -7.685  -1.801   5.533  59.921
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  135.315      6.871   19.69   <2e-16 ***
## xGDPPCP      -12.231      0.723  -16.92   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.64 on 205 degrees of freedom
## Multiple R-squared:  0.5827, Adjusted R-squared:  0.5806
## F-statistic: 286.2 on 1 and 205 DF,  p-value: < 2.2e-16
```

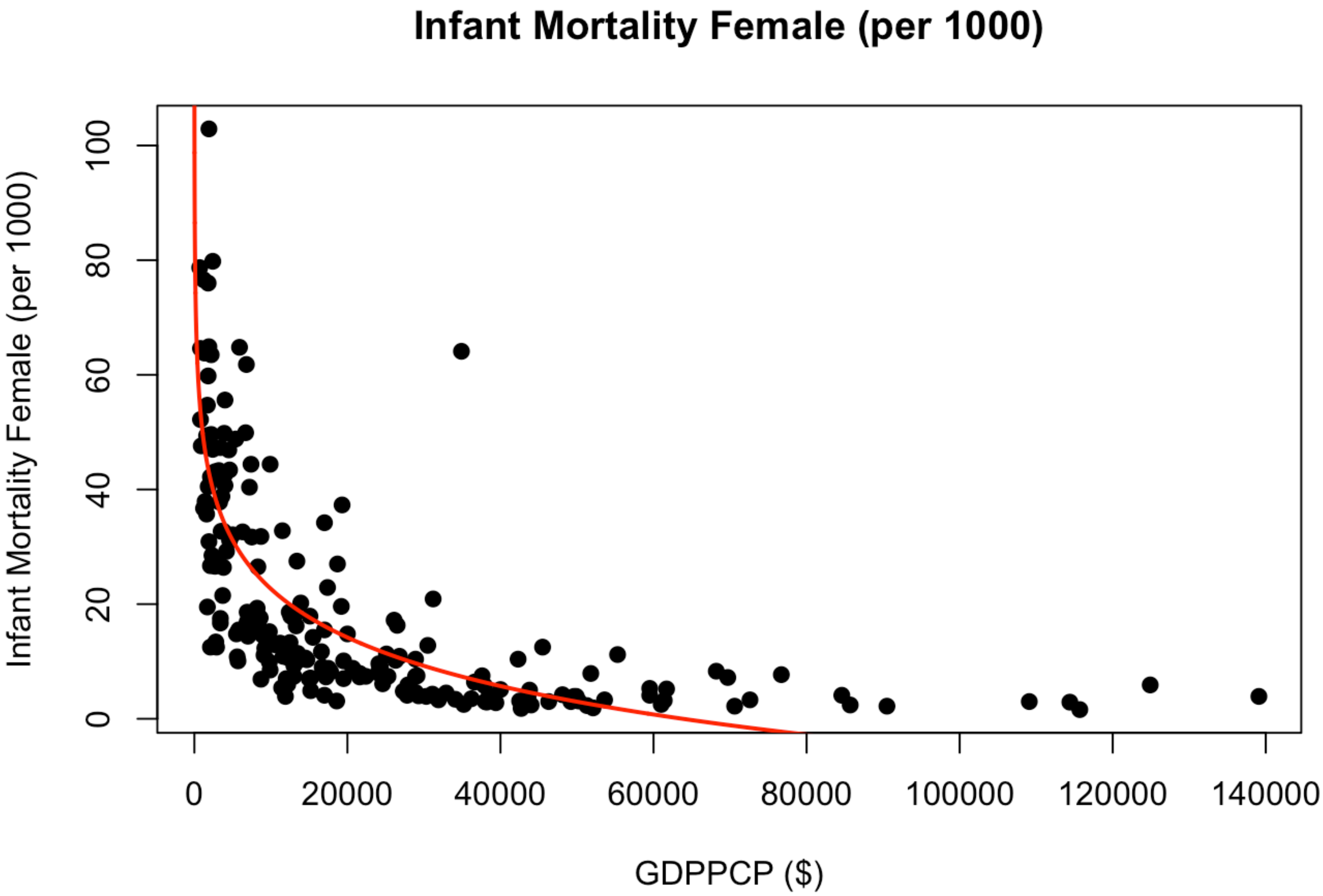
```
r_sqr <-summary(fit_MU5F)$r.squared
r_sqr <- round(r_sqr,digits=3)

plot(xGDPPCP, yMU5F, main="Life Expectancy at Birth Female", xlab="GDPPCP ($)", yl
ab="Life Expectancy at Birth Female", pch=19)
abline(fit_MU5F, col="blue") # regression line (y~x)
text(80,3.75,"R Squared is:")
text(80,3.5,r_sqr)
```




```
# LIFE EXPECTANCY REGRESSION CURVE MALE
x  <- seq(0,90,0.01)
x_hat <- exp(x)
y <- -14.8730*x+164.6484
y_hat <- exp(y)
plot(CIAData$GDPPCP, CIAData$MU5M, main="Infant Mortality Male (per 1000)", xlab="
GDPPCP ($)", ylab="Infant Mortality Male (per 1000)", pch=19)
points(x_hat,y, type='l', col="red", lwd = 2)
```




```
# LIFE EXPECTANCY REGRESSION CURVE FEMALE
x  <- seq(0,90,0.01)
x_hat <- exp(x)
y <- -12.231*x+135.315
y_hat <- exp(y)
plot(CIAData$GDPPCP, CIAData$MU5F, main="Infant Mortality Female (per 1000)", xlab
="GDPPCP ($)", ylab="Infant Mortality Female (per 1000)", pch=19)
points(x_hat,y, type='l', col="red", lwd = 2)
```


Infant Mortality Female (per 1000)
Infant Mortality Female (per 1000)
0
20
40
60
80
100
0
20000
40000
60000
80000
100000
120000
140000
GDPPCP ($)

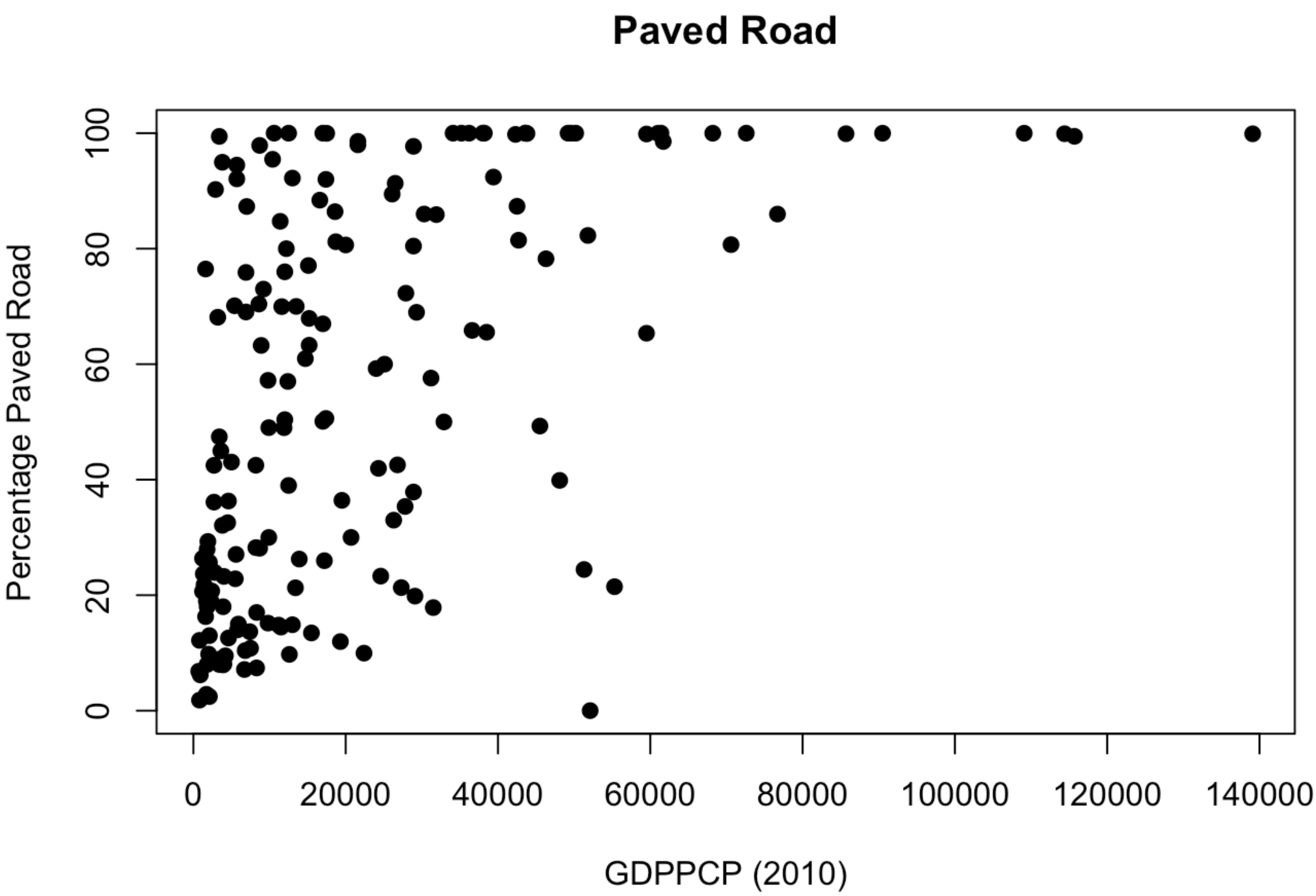
B.5. Percentage of Paved Roads

Notebook Thesis Rémon ten Bhömer TU Delft 2018: Regression Curve Percentage of Paved Roads

```
library("readxl")
RoadData <- read_excel("RegressiePaved.xlsx", 1)
```

```
df <- data.frame(RoadData$country,
                 RoadData$`gdp per capita purchasing power parity usd`,
                 RoadData$`Percentage of Paved Roadway`)

colnames(df) <- c("Country",
                  "GDPPCP",
                  "RoadPaved")
```

```
# EXAMINE ROAD DATA
x  <- seq(0, 1500000, 1)
plot(df$GDPPCP, df$RoadPaved, main="Paved Road", xlab="GDPPCP (2010)", ylab="Perce
ntage Paved Road", pch=19)
```

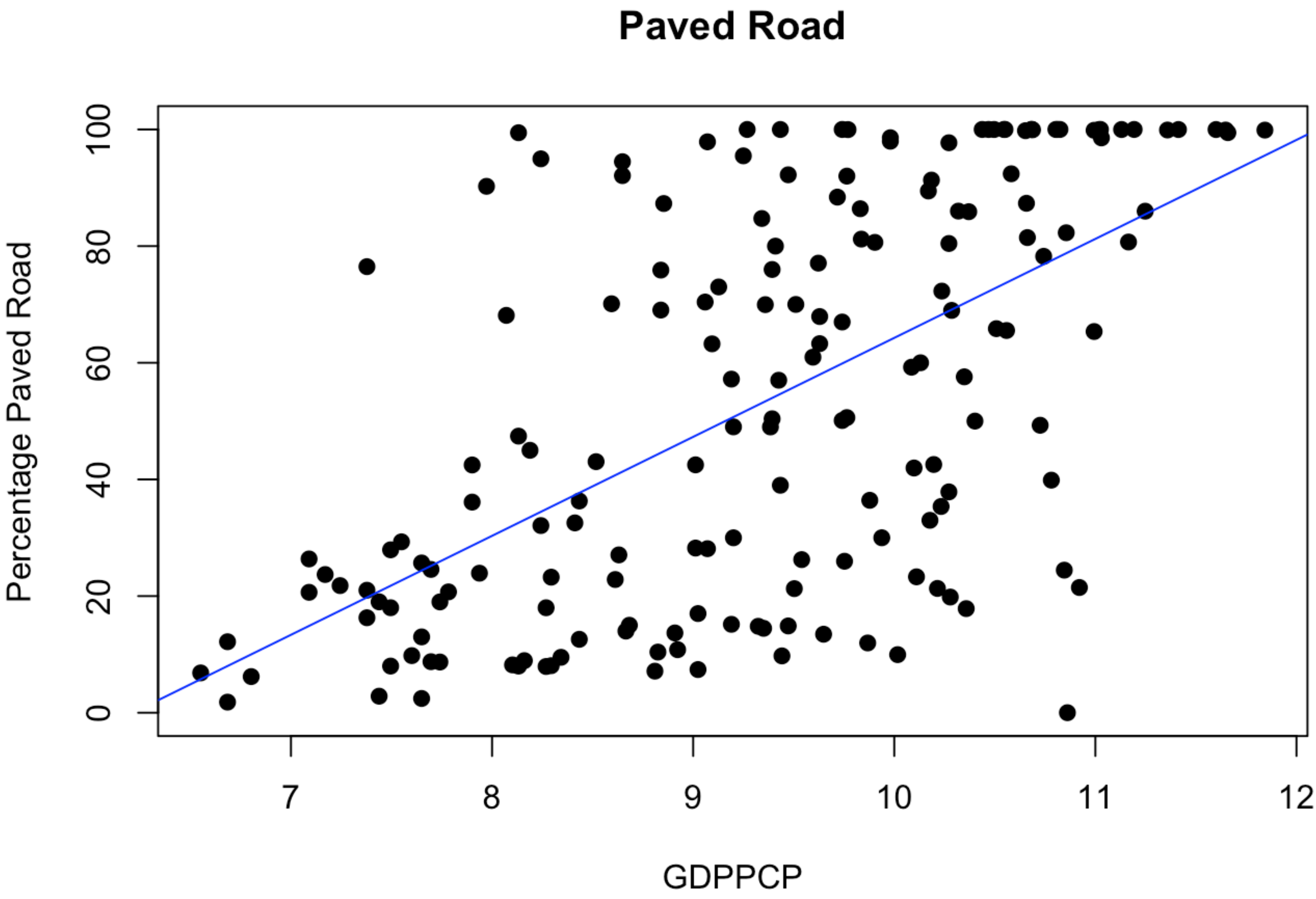


```
# TRANSFORM DATA FOR PAVED ROADS
# Here I choose various data transformations to fit the data
# Here I take the log of the GDPPCP (2010) because there are such wide variations
in the dependent variable.
RoadData$link_GDPPCP <- log(df$GDPPCP)
# This is just a linear relationship; no transformation
RoadData$link_Road <- df$RoadPaved

# MODEL THE DATA
# Rename the variables with dependent and independent coding
yRoad <- RoadData$link_Road
xGDPPCP <- RoadData$link_GDPPCP
```

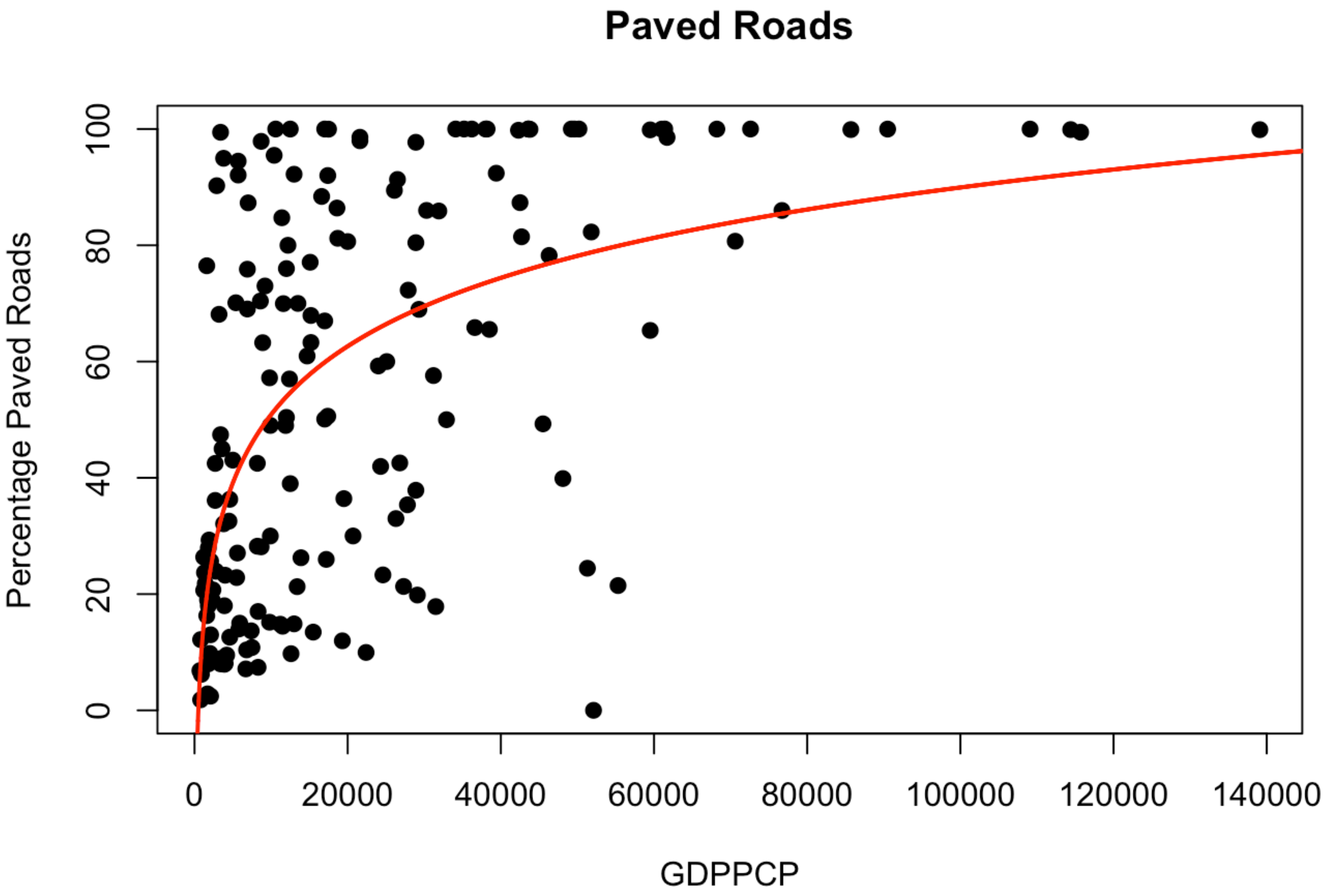
```
# REGRESSION ANALYSIS FOR PAVED ROADS
fit_Road <- lm(yRoad~xGDPPCP)
summary(fit_Road) # show results
```

```
##
## Call:
## lm(formula = yRoad ~ xGDPPCP)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -78.872 -20.795   1.337  19.342  66.878
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -105.431     15.975   -6.60 4.79e-10 ***
## xGDPPCP       16.970      1.692   10.03  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 27.45 on 174 degrees of freedom
## Multiple R-squared:  0.3662, Adjusted R-squared:  0.3626
## F-statistic: 100.6 on 1 and 174 DF,  p-value: < 2.2e-16
```

```
r_sqr <-summary(fit_Road)$r.squared
r_sqr <- round(r_sqr,digits=3)

plot(xGDPPCP, yRoad, main="Paved Road", xlab="GDPPCP", ylab="Percentage Paved Road
", pch=19)
abline(fit_Road, col="blue") # regression line (y~x)
text(80,3.75,"R Squared is:")
text(80,3.5,r_sqr)
```




```
# REGRESSION CURVE PERCENTAGE OF PAVED ROADS
x  <- seq(0, 1500000, 1)
plot(df$GDPPCP, df$RoadPaved, main="Paved Roads", xlab="GDPPCP", ylab="Percentage
Paved Roads", pch=19)
lines((16.970*log(x))-105.431, col="red", lwd = 2)
```


Paved Roads
Percentage Paved Roads
0
20
40
60
80
100
0
20000
40000
60000
80000
100000
120000
140000
GDPPCP

C

EMA Open Exploration Script

EMA Open Exploration

Written by: Dr. Erik Pruyt and Stefan Wigman

Used by: R. ten Bhömer as part of masters thesis

This script is used to perform an open exploration on the Geopolitical model. 500 scenario runs have been performed. This analysis has been performed for the Unites States

1. Importing the required Python packages

In [ ]:

```
from ema_workbench import(Model, RealParameter,Constant,IntegerParameter,Categ
oricalParameter, TimeSeriesOutcome,\
                          Policy, perform_experiments, ema_logging, save_resul
ts, load_results)
from ema_workbench.connectors.vensim import VensimModel
from ema_workbench.em_framework.evaluators import LHS, SOBOL
import timeit
from ema_workbench import MultiprocessingEvaluator

ema_logging.log_to_stderr(ema_logging.INFO)

import numpy.random
numpy.random.seed(123456789)
```

2. Loading the Vensim model

In this step, we specify the working directory (this should be the location of this notebook and the Vensim model file), and import the Vensim model into our workspace. Note that the model file needs to have a .vpm extension, which can be created by using the publish function in Vensim.

In addition, the Vensim dll can be specified if required. This defaults to 'vendll32.dll' for the normal vensimdll, but if double precision is installed and required, change this to 'vdpdll32'. [Note: Since only one model can be loaded per dll file, this also allows for two models to be loaded at the same time.]

We can also specify a name for the .vdf file in which we want Venpy to store the intermediate results. Note that this file should not be in use by any other program. This defaults to 'CurrentRun'.

In [ ]:

```
wd =r'C:\Users\Rémon ten Bhömer\Documents\EMA\Models'
model = VensimModel('ThesisFinalModelTenBhomer', wd = wd , model_file='ThesisF
inalModelTenBhomerVPM.vpm')
ema_logging.log_to_stderr(ema_logging.INFO)   # we want to see what EMA is doi
ng
```

3. Specify uncertainties and outcomes

Here we will specify the uncertainties and outcomes. The subscript of a variable is simply placed between square brackets.

Every range between 0 and 1 is supposed to be a percentage. If other values are assigned they will be explained in the script.

In [ ]:

```
uncertainties = [
            RealParameter('Influence Factor Corruption on Autocracy Democracy[
United States]', 0, 1),
#Judiciary Change Excelerator is a growth or decline rate with -1% or 1%
            RealParameter('Judiciary Change Excelerator[United States]', -0.01
, 0.01),
            RealParameter('Influence Factor Corruption[United States]', 0, 1),
            RealParameter('Influence Factor Judicidary[United States]', 0, 1),
            RealParameter('Influence Factor Economic Openess[United States]',
0, 1),
            RealParameter('Influence of Ethnicity Size Change[United States]',
0, 1),
            RealParameter('Influence of Institutional Instability[United State
s]', 0, 1),
#Conflict
            RealParameter('Influence Factor Grievances[United States]', 0, 1),
            RealParameter('Influence Factor Discrimination[United States]', 0,
1),
            RealParameter('Influence Factor Reaction[United States]', 0, 1),

#Ethnicity Growth
            RealParameter('Percentage of Total Population Growth to First Ethn
icity[United States]', 0, 1),
            RealParameter('Percentage of Total Population Growth to Second Eth
nicity[United States]', 0, 1),
            RealParameter('Percentage of Total Population Growth to Third Ethn
icity[United States]', 0, 1),
            RealParameter('Percentage of Total Population Growth to Other Ethn
icities[United States]', 0, 1),
            RealParameter('Percentage of Total Population Growth to Fourth Eth
nicity[United States]', 0, 1),

#Uncertainties for Power Sum of Total State Power
            RealParameter('Influence Factor Total Population[United States]',
0, 1),
            RealParameter('Influence Factor Land Mass[United States]', 0, 1),
```

```
            RealParameter('Influence Factor GDP per Capita[United States]', 0,
1),
            RealParameter('Influence Factor Relative Army Strength[United Stat
es]', 0, 1),
            RealParameter('Influence Factor Autocracy Democracy[United States]
', 0, 1),

#Uncertainties for Power Sum of Military
            RealParameter('Influence factor Airforce on Military Power[United
States]', 0, 1),
            RealParameter('Influence Factor Army on Military Power[United Stat
es]', 0, 1),
            RealParameter('Influence Factor Navy on Military Power[United Stat
es]', 0, 1),
            RealParameter('Influence Factor Active Personel on Military Power[
United States]', 0, 1),
            RealParameter('Influence Factor Mean Years of Education on Militar
y Power[United States]', 0, 1),

#Resources
            RealParameter('Influence factor Oil Price[United States]', 0, 1)
                 ]

outcomes = [

    TimeSeriesOutcome('State Power[United States]'),
    TimeSeriesOutcome('Autocracy Democracy[United States]')

           ]
```

In [ ]:

```
model.uncertainties = uncertainties
model.outcomes = outcomes
```

In [ ]:

```
nr_scenarios =500
```

In [ ]:

```
# policy_results = perform_experiments(model,scenarios=nr_scenarios,policies=p
olicies)
```

In [ ]:

```
start_time = timeit.default_timer()

with MultiprocessingEvaluator(model) as evaluator:
    policy_results = evaluator.perform_experiments(scenarios=nr_scenarios)

elapsed = timeit.default_timer() - start_time
```

In [ ]:

```
save_results(policy_results, r'C:\Users\Rémon ten Bhömer\Documents\EMA\Results
\RemonEMAOpenExplorationUnitedStates500.tar.gz')
```

In [ ]:

```
# VERGELIJKEN MET... save_results(policy_results, r'C:\Users\LocalAdmin\Docume
nts\AC_Project\SC00SC00SC14runs500v71delayorder23.tar.gz')
```

5. Slicing

6. Visualization of the results

In [ ]:

```
from ema_workbench import(Model, RealParameter,Constant,IntegerParameter,Categ
oricalParameter, TimeSeriesOutcome, Policy, perform_experiments, ema_logging,
save_results, load_results)
# from ema_workbench.connectors.vensim import VensimModel
# from ema_workbench.em_framework.evaluators import LHS, SOBOL
# import timeit
# from ema_workbench import MultiprocessingEvaluator

## Import specific plotting commands:
import matplotlib.pyplot as plt
from ema_workbench.analysis.plotting import lines, plot_lines_with_envelopes,
envelopes
from ema_workbench.analysis.plotting_util import KDE, HIST, VIOLIN, BOXPLOT
import seaborn as sns
import ema_workbench.analysis.pairs_plotting as pairs
import ema_workbench.analysis.plotting as emaplt
```

In [ ]:

```
%matplotlib inline
import seaborn as sns
sns.set()
plt.rcParams["figure.figsize"] = (20,8)
```

In [ ]:

```
# policy_results = load_results(r'c:\users\wvanderpaauw\Work Folders\Documents
\Thesis\Wijnand.tar.gz')
```

Visualization Results

In [ ]:

```
policy_results = load_results(r'C:\Users\Rémon ten Bhömer\Documents\EMA\Result
s\RemonEMAOpenExplorationUnitedStates500.tar.gz')
```

In [ ]:

```
# policy_results = sliced_results
```

In [ ]:

```
experiments, outcomes = policy_results
```

In [ ]:

```
figure = lines(policy_results, density = 'kde')
plt.show()
```

D

Web Scraper Global Fire Power Index

Web Scraper Global Fire Power Index

Written by: Rémon ten Bhömer

Year: 2018

This web scraper is built to scrape data for all countries from https://www.globalfirepower.com (https://www.globalfirepower.com) to a comma seperated csv.

Imports

In [ ]:

```
import pandas as pd
import numpy as np
import matplotlib
import seaborn
import urllib
from bs4 import BeautifulSoup
import re
import ssl
from urllib.request import urlopen as uReq
from urllib.request import Request, urlopen
import requests
import csv
```

Load the web page with the current ranking

In [ ]:

```
source = requests.get('https://www.globalfirepower.com/countries-listing.asp')
.text
soup = BeautifulSoup(source, 'lxml')
```

Within this website search for everything with 'country-military-strength-detail' and add the standard web adress to complete the urls for all countries

In [ ]:

```
Country = soup.find_all("a")

country_list_all = []

for link in Country:
    country_list = str('https://www.globalfirepower.com'+link.get('href'))
    if 'country-military-strength-detail' in country_list:
        country_list_all.append(country_list)
```

Get list of country names from the earlier created list of urls

In [ ]:

```
def remove_cruft(s):
    return s[80:]

country_names = [remove_cruft(s) for s in country_list_all]
```

Load the first web page (in this case the United States of America). It is nessecary to process one country before being able to iterate for all countries

In [ ]:

```
source = requests.get('https://www.globalfirepower.com/country-military-streng
th-detail.asp?country_id=united-states-of-america').text
soup = BeautifulSoup(source, 'lxml')
```

Search for certain text sizes and colours

In [ ]:

```
my_list = soup.find_all('span', class_='textLarger textBold textWhite')
my_list2 = soup.find_all('span', class_='textLarger textBold')
```

In [ ]:

```
my_list3 = my_list + my_list2
```

Create empty list and for each item in the list get the number out of and then get the '.' out of the numer and make it a float. It is important to add the country in front of the list because otherwise you can't make a dataframe with a country column.

In [ ]:

```
num_list_final = []
for item in my_list3:
    numbers = re.compile('\d+(?:\.\d+)?')
    num = numbers.findall(str(item))
    num = ''.join(num)
    try:
        num = float(num)
    except:
        num = 0
    num_list_final.append(num)

num_list_final.insert(0, 'USA')
```

Create dataframe with column names

In [ ]:

```
df = pd.DataFrame([num_list_final])
df.columns = ['country','total Population', 'Total Aircraft Strenth', 'Total Navy Assets', 'Consumption (bbl/dy)', 'External Debt ($)','Coastline', 'Shared Borders', 'Manpower Available', 'Fir for Service', 'Reaching Military Age', 'Active Personel', 'Reserve Personnel', 'Fighter Aircraft', 'Attack Aircraft', 'Transport Aircraft', 'Trainer Aircraft', 'Total Helicopter Strength', 'Attack Helicopter', 'Combat Tanks', 'Armored Fighter Vehicles', 'Self Propelled Artillery', 'Towed Artillery', 'Rocket Projectors', 'Aircraft Carriers', 'Frigates', 'Destroyers', 'Corvettes', 'Submarines','Patrol Craft','Mine Warfare Vessels', 'Production (bbl/dy)', 'Proven Reserves (bbl)', 'Labor Force', 'Merchant Marine Strength', 'Major Ports/Terminals', 'Roadway Coverage (km)', 'Railway Coverage (km)', 'Servicable Airports', 'Defense Budget', 'Foreign Exchange/Gold', 'Purchasing Power Parity', 'Square Land Area', 'Waterways']
```

Iterate over the country list to get all values

In [ ]:

```
for i, x in enumerate(country_list_all):
    source = requests.get(x).text
    soup = BeautifulSoup(source, 'lxml')

    mylist = soup.find_all('span', class_='textLarger textBold textWhite')
    mylist2 = soup.find_all('span', class_='textLarger textBold')

    mylistfinal = mylist + mylist2

    num_list_final = []
    for item in mylistfinal:
        numbers = re.compile('\d+(?:\.\d+)?')
        num = numbers.findall(str(item))
        num = ''.join(num)
        try:
            num = float(num)
        except:
            num = 0
        num_list_final.append(num)

    country_name = country_names[i]

    num_list_final.insert(0, country_name)

    series = pd.Series(num_list_final, index=df.columns)
    df.loc[i] = series
```

In [ ]:

```
df
```

Write data frame to CSV

In [ ]:

```
df.to_csv('MilitaryCapabilityData.csv')
```

For now there was one value which I left out because I found it unnessecary. It can be added by adding the below line of code to the list sum

In [ ]:

```
mylist3 = soup.find_all('span', class_='textLarger textBold textDkBlue')
```